CUK-H06043-23-P3542
370.1505
B777

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

23

JUNE 1984

VOL. 54, PART 2

THE
BRITISH JOURNAL
OF
EDUCATIONAL PSYCHOLOGY

Edited by
Professor GEOFFREY BROWN

Assistant Editors:
Dr. R. L. BURDEN
Dr. C. W. DESFORGES
Dr. D. FONTANA

Sub-Editor:
Dr. S. CUNNINGHAM

Book Review Editor:
Dr. C. W. DESFORGES

Advised by the Editorial Committee:
Dr. H. BELOFF
Professor M. B. SUTHERLAND
Professor J. WRIGLEY

P35412

Vol. [illegible] Part 2

ish Jou

The British Journal [illegible] gy is issued by the British Psychological Society once a term, in Fe[illegible] ation November. The subscription is £15·00 a year post free, payable in adva[illegible] number.

Subscriptions and Ord [illegible] sent to Scottish Academic Press, 33 Montgomery Street, Edinburgh, EH7 [illegible] Members of the British Psychological Society receive the *Journal* on special terms.

Papers for publication in duplicate should be sent to Professor Geoffrey Brown, School of Education, University of East Anglia, Norwich NR4 7TJ. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for *The British Journal of Educational Psychology* should be sent to: Dr. C. W. Desforges, School of Education, University of East Anglia, Norwich NR4 7TJ.

Advertisements for this *Journal* should be sent to Scottish Academic Press.

Published by
SCOTTISH ACADEMIC PRESS, 33 MONTGOMERY STREET, EDINBURGH, EH7 5JX,
for the British Journal of Educational Psychology, Ltd.

All rights reserved

© 1984. The British Journal of Educational Psychology

PRINTED IN GREAT BRITAIN BY CLARK CONSTABLE, EDINBURGH, LONDON, MELBOURNE

370.1505
B777

Br. J. educ. Psychol., 54, 137-151, 1984

23

A HIERARCHY OF INTELLECTUAL DEVELOPMENT IN SCIENCE

By R. G. HACKER
(*Department of Education, The University of Western Australia*)

SUMMARY. Three schools of thought on the intellectual development of science pupils were examined and it was concluded that theories which take account of the characteristics of both the subject-matter and the learner have the best prognosis. Strengths and weaknesses of popular methodologies for eliciting information about cognitive development were explored and it was suggested that pupils' classroom behaviours can provide a useful data-base for the cognitive theorist. These points were illustrated by combining a logical analysis of a scheme of generic categories of scientific behaviours with a study of the incidences of these behaviours in science lessons across twelve consecutive years of primary and secondary schooling. Behaviours of 3,751 pupils in 864 lessons taught by 144 science teachers were cluster-analysed and a three-tier, hierarchical model of intellectual development in science was postulated. The preferred instructional strategies of the science teaching population studied were described.

INTRODUCTION

WHEREVER theoretical uncertainty exists, it is inevitable that conflicting schools of thought will emerge. Lovell (1980) has characterised two such schools of thought on the intellectual development of children. Within one school, the logical aspects of learning are emphasised. Intellectual development is viewed primarily as the process of building in children's minds the intellectual structures associated with a discipline. The possibilities of age-related restrictions on the types of thought-processes in which children can engage are disregarded and thinking is seen as being only *quantitatively* different at different ages. The theoretical positions associated with Ausubel (1968), Gagné (1977) and Trabasso (1977) are exemplars of this viewpoint.

A second school of thought emphasises the psychological aspects of the learning process. Here the theoretical focus is on internal restrictions on the child's thinking and *qualitative* changes in this thinking as cognitive growth occurs. Differences in the intellectual structures associated with different disciplines and fields of study are ignored and an attempt is made to build 'content-free' intellectual structures which transcend discipline boundaries. This school of thought encompasses the Piagetian (Inhelder and Piaget, 1958) and neo-Piagetian paradigms and information-processing views of cognitive development (see, for example, Klahr and Wallace, 1976).

It is true that there are epistemological and other differences between theorists within each of these broad schools of thought. Nonetheless, Lovell (1980) points out that these differences are of a lower order of magnitude than the between-group differences which have been outlined.

These theoretical perspectives have quite different practical implications for teaching and learning in science. Logical theorists seek to focus attention on some form of task analysis which makes the subject-matter more digestible for the learner, Gagné's work providing a well-known example of this approach. A key element in Gagné's approach is the development of a student learning hierarchy. This is achieved by placing a target intellectual ability at the apex of a pyramid and then undertaking a logical analysis of the topic to identify prerequisite intellectual abilities and thus fill in the body of the pyramid. The science teacher must then develop diagnostic pre-tests and post-tests for each intellectual ability incorporated into the pyramid and performances on these tests reveal the full and partial intellectual structures acquired by the learner. In this way, diagnostic tests determine precisely what has to be learned to work towards the apex of the pyramid.

As Case (1975) points out, this theory takes no account of the possibility of age-related restrictions on the types of intellectual abilities which children can acquire. It is possible that Gagné's earlier work with USAF pilot training programmes may go some way towards accounting for his perspective on learning. Working with able, mature adult minds would naturally focus attention on the need for a thorough task analysis as the major limiting factor in the learning process. Certainly, the bulk of the research work based on logical theories of intellectual development has been carried out with senior secondary or tertiary level students, such students being regarded by psychological theorists as having completed their basic intellectual development.

By way of contrast, psychological theorists see the internal cognitive processing capacities of pupils as the main focus for teaching and learning in science. Subject-matter is important only in so far as it must be matched to these capacities and any analysis of this subject-matter must be approached from the child's perspective rather than that of the discipline specialist.

Piaget's work provides an exemplar for this viewpoint which has been particularly influential in the field of science education. Children are thought to assimilate knowledge from their environment and pass through four major developmental stages in an invariant sequence, each stage being characterised by the emergence of new kinds of intellectual abilities. The science teacher's task is firstly to assess the level of development of the learner and then to select and sequence learning experiences in such a way as to facilitate progress through these stages (rather than according to a logical sequence inherent in the subject-matter). Shayer and Adey (1981) provide a useful articulation of this position for science educators.

Piaget's later work (Piaget, 1971) reveals fully the extent to which his initial training as a biologist influenced his theory of intellectual development. He regarded intelligence as a specialised form of biological functioning and saw striking parallels between the physical growth of the embryo and the mental growth of children. He believed that the study of intellectual development would eventually become firmly established as a branch of biology.

It is interesting to note Shayer's (1980) suggestion that one possible reason for the influence of Piaget's theory in the field of science education is that it may turn out to be merely a description of the growth of a sub-set of *scientific* thinking abilities rather than a complete metatheory. Certainly, at the level where the theory is articulated as descriptions of developmental stages, both the administered tasks and the inferred intellectual abilities are firmly located within the discipline boundaries of the physical sciences.

The flow of Piagetian literature and the enormous impact on science curriculum development certainly attest to the appeal of Piaget's theory. However, Flavell (1963), Brainerd (1978) and Brown and Desforges (1979) draw attention to some of the limitations of Piaget's model and research methodology. Whereas the theory is both powerful and parsimonious, in many respects it is untestable. Even where the theory is expressed in testable terms, it turns out to be a poor predictor of pupil behaviours (Lawson *et al.*, 1978).

Of course, this logical *vs.* psychological polarisation amongst those with an interest in the learning process is not new. As early as the turn of the century, Dewey (1902) had forewarned us of the divisions which were to emerge amongst cognitive theorists when he observed:

> " The fundamental factors in the educative process are an immature, undeveloped being; and certain social aims, meanings, values incarnate in the mature experience of the adult. The educative process is the due interaction of these

> forces. Such a conception of each in relation to the other as facilitates completest and freest interaction is the essence of educational theory.
>
> But here comes the effort of thought. It is easier to see the conditions in their separateness, to insist upon one at the expense of the other, to make antagonists of them, than to discover a reality to which each belongs. . . . When this happens a really serious practical problem—that of interaction—is transformed into an unreal, hence insoluble, theoretic problem."

Although Dewey had put us clearly in mind of the limitations of either a purely logical or purely psychological approach to studying learning, perhaps Bruner (1960) and Schwab (1964) were the first major cognitive theorists to begin to recognise the importance of both the developing child and the structure of a discipline as factors to be considered in studying the learning process. It is clear that these theorists provided the point of departure for a third school of thought which Lovell (1980) has chosen to disregard.

More recently, there have been a number of promising developments within this third school. A number of trends can be discerned. A broader range of theoretical constructs is being employed than previously; many theorists are now focusing their attention on more specific tasks and topics, believing that the time is not ripe for an all-embracing metatheory of intellectual development; others are attempting to build bridges between existing theoretical positions.

This third school of thought on the intellectual development of science pupils subsumes research studies concerned with children's alternative frameworks (see, for example, Driver, 1982; Osborne and Bell, 1983). These studies emphasise the importance of the ideas which children bring to the classroom and the content and context of a task as factors to be considered in studying the learning process.

The work of Fensham *et al.* (1981) provides an illustration of this type of study. They have carried out logical analyses of the intellectual structures of specific science curriculum topics in terms of ' intended cognitions ' and have then used a variety of psychological interview techniques to probe the intellectual structures possessed by individual students, thus assessing the extent to which they match the ' intended cognitions '.

Recent years have also witnessed a broadening of the range of methodologies used by cognitive theorists for eliciting information about pupils' cognitive development. The ubiquitous, though time-consuming, clinical method for probing cognitive structures has been subjected to increasing criticism (Flavell, 1963; Brainerd, 1978). Brown and Desforges (1979) conclude that it provides a poor foundation upon which to rest a theory. The price paid for flexibility in following children's responses to variants of a problem-solving situation is an experiment which cannot be replicated. The bulk of the data collected is in the form of subjects' protocols and reporting is inevitably selective. The impossibility of genuine replications means that observer objectivity and reliability cannot be demonstrated. Perhaps the methodology is better perceived as an ethnographic technique than as a psychometric method of collecting information.

It is suggested here that pupils' classroom behaviours can provide a rich, though often neglected, data-base for the cognitive theorist, providing steps are taken to ensure that information is collected in a way which ensures that it is both reliable and representative. Natural behaviour sequences avoid the artificiality of tests administered by research workers and associated questions about their validity. Also, in the classroom, whole sequences of behaviours frequently focus on development of a specific intellectual ability, thus providing cross-checks on the inferences drawn from pupil behaviours of a kind not available with isolated test items.

It seems reasonable to expect that the profiles of abilities which science pupils will ultimately develop must reflect closely the profiles of abilities which they are able to practise in the classroom. If, for example, science pupils were never to participate in designing experiments during their science lessons, then it would seem unlikely that the requisite abilities would be developed. The position taken here is that intellectual growth in science is largely determined by the opportunities provided *in the classroom* for pupils to practise the intellectual skills commonly associated with science and scientific inquiry.

Studies such as those by Fensham *et al.* (1981), carried out at the level of specific science topics and tasks, have the potential to provide valuable information for the teacher and curriculum developer by suggesting optimal intra-topic or intra-task learning sequences. However, there is also a clear need for studies which attempt to provide a more general overview of intellectual development in science and thus guide inter-topic sequencing and course development over a number of years of schooling. The study reported here aims to take account of the structure of science and the characteristics of the developing child at this more general level. Firstly, as a logical component of the study, a scheme of generic categories of pupil behaviour will be developed, the occurrence of which might be regarded as evidence for the attainment of the intellectual structures associated with science. Then, the psychological component will be provided by a mapping of the development of these scientific thinking skills via an observational study of the incidences of these behaviours in science lessons across the primary and secondary school.

The purpose of the study is to provide a description of the profiles of intellectual abilities actually practised by pupils in science lessons and to investigate changes in these profiles as a function of pupil age. It is hoped that this will result in a better understanding of the steps that intervene between children's initial experiences of science and their eventual assimilation of the intellectual structures possessed by adult scientists.

METHOD

Instrument

The Science Lesson Analysis System (SLAS) has been described fully elsewhere (Hacker, 1982). A rationale for the categories incorporated into the instrument was provided and each category was delimited by way of an operational definition and instances of specific behaviours. A users' manual, observer training programmes and observer reliability assessment procedures were also provided.

The SLAS instrument is illustrated in Figure 1. Each behaviour to be recorded is classified according to both its *form* and its *function*. The columns of the instrument provide nine categories for the form of the interaction and the rows of the instrument provide 12 categories for interactive functioning, i.e., the intellectual ability which the pupil practises as a result of the interaction.

The categories for the function of the interaction reflect Schwab's (1960, 1962, 1964) scholarly analysis of the scientific enterprise and include categories of behaviour commonly associated with the substantive domain of science and the syntactic domain, including both stable and fluid aspects. They might be regarded as a condensed version of Klopfer's (1971) behaviours or as a version of the categories of Eggleston *et al.* (1975), expanded to include behaviours likely to occur in primary school science lessons.

In the context of these categories, problem-solving is defined as a convergent activity where only one correct solution is perceived by the pupil as being acceptable to the teacher. For all interactions, whether teacher or pupil-initiated, it is the intellectual demand actually placed on the pupil, rather than the teacher's intent, which forms the basis for judging the function of the interaction.

FIGURE 1

THE SCIENCE LESSON ANALYSIS SYSTEM

THE SCIENCE LESSON ANALYSIS SYSTEM

	VERBAL INTERACTIONS					NON-VERBAL INTERACTIONS With Resource Materials			
	TEACHER — INITIATED			STUDENT — INITIATED		SCIENCE MATERIALS		MULTI-MEDIA MATERIALS	
	A	B	C	D	E	F	G	H	I
THE INTELLECTUAL ABILITY BEING PRACTISED:	QUESTIONS answered by	STATEMENT concerned with	DIRECTIVE for the purpose of	REFERRAL to teacher for the purpose of.	CONSULTATION with student for the purpose of	TEACHER interacts for the purpose of	STUDENT interacts for the purpose of:	TEACHER interacts for the purpose of.	STUDENT interacts for the purpose of
1. Acquiring, recalling or confirming facts									
2. Delineating scientific concepts, principles or theoretical models									
3 Identifying problems									
4. Solving concrete problems									
5. Solving problems by applying scientific concepts, principles or models									
6. Making or testing hypothesis or speculation									
7. Identifying or describing apparatus, equipment or materials									
8 Describing or practising conventional experimental procedures									
9. Designing novel experimental procedures									
10. Making, describing or recording observations									
11. Interpreting observed or recorded data									
12 Inferring from observed or recorded data									

When the form and function of the interaction have been determined, the appropriate cell of the instrument's matrix is selected. A sign system for recording information is utilised whereby each cell is marked present or absent during a three-minute period in the classroom. This time period was empirically determined to reflect a compromise between the need for the most detailed record possible and the limitations on the speed with which a trained observer is able to perform reliably the sometimes complex tasks of observing and coding interactions (see also Eggleston and Galton, 1981). One copy of the SLAS matrix is completed for each three-minute period of observation.

For a particular classroom, these raw data are summarised over the total period of observation by taking each row and column of the instrument in turn and calculating the proportion of the matrices with an entry in that row or column. This procedure results in the observational data for each classroom being summarised as a profile of 21 probabilities: a sub-set of 9 relating to the likelihood of interactions of various forms occurring in that classroom and a sub-set of 12 relating to the frequencies with which pupils engage in the intellectual behaviours incorporated into the instrument.

Sample

Data were collected from 144 teacher-class units located in 62 state primary and secondary schools in Western Australia, this sample consisting of 12 teacher-class units at each of the 12 consecutive years of primary and secondary schooling. Science was both substantially and discretely represented in the school curriculum at each year level and a carefully co-ordinated and sequenced programme covered all 12 years of primary and secondary schooling.

Children entered primary school at the start of the year during which they reached 6 years of age and throughout the primary school (years 1 to 7) an integrated programme in the biological and physical sciences was offered at each year level. At the lower secondary level (years 8 to 10), a general science curriculum was studied, comprising discrete biology, chemistry and physics topics. At the upper secondary level (years 11 and 12) these disciplines were formally differentiated.

Earlier studies (Eggleston *et al.*, 1976; Hacker *et al.*, 1979; Hacker, 1980) had shown that the discipline area being studied would influence classroom interactions. Therefore, at each secondary year level, equal numbers of classes studying biology, chemistry and physics topics were selected to provide a stratified, purposive sample with respect to the discipline area being studied.

Classes were generally unstreamed but in schools where streaming was in evidence at the higher year levels, a class from the middle ability range in science was selected. Teachers with less than two years of teaching experience were excluded from the study, as were schools with facilities considered to be so poor as to preclude the usual range of science teaching and learning strategies at that level.

Participation rates for teachers were in excess of 90 per cent for the secondary science teachers approached and approximately 70 per cent for primary teachers. The most common reason given for the non-participation of primary teachers was that their science teaching was delegated to a specialist teacher in the school or, in some cases, to an itinerant specialist travelling between schools. The distribution of number of years of experience for the teachers was positively skewed with a mean value of 9·1 years. The proportion of male teachers in the sample increased almost linearly from 0·08 at the lower primary level (years 1 and 2) to roughly 0·75 at the upper secondary level (years 11 and 12). Mean class sizes were 28 at the primary school level and 24 at the secondary school level.

Procedure

The three observers selected were science graduates who were also qualified science teachers with substantial teaching experience. The classroom observation work was preceded by one full week of intensive observer training and at the end of this week observer reliability trials were carried out. Throughout the data collection phase of the study, further reliability checks and retraining sessions were held as required to maintain reliability levels.

The principals of the schools selected provided access to teachers who were asked if they were willing to participate in the project. These teachers were asked to choose six lessons which they considered to be typical of their teaching patterns. Visits to classrooms by different observers were randomised to avoid systematic observer errors in the results.

Steps were taken to control the influences of variables extraneous to the research design and thus ensure that the data collected were representative of the patterns of behaviours at each year level. Observers were instructed to discard data from lessons where factors such as ineffectual class management, lack of rapport with the class or the teacher allowing the lesson to be dominated by an intellectually precocious minority were considered to have distorted the patterns of behaviours.

In schools where the presence of an observer in the classroom was a novelty, data were discarded from the initial lessons observed to allow time for the class to accommodate to the new situation. Observers were also instructed to ensure that the recorded data reflected the teacher's normal usage of facilities and resources. If, for example, the teacher's normal teaching pattern involved a double lesson in a laboratory followed by a single lesson in the classroom, then this ratio was reflected in the allocation of observation time. During class experiments, small-group discussions and periods of project work, observers were required to move between groups of students to ensure the representativeness of the information recorded for the whole class.

In addition to the information coded on the SLAS instrument, observers were required to compile an anecdotal abstract for each lesson which identified the *specific* intellectual abilities developed during the lesson, the subject matter of the lesson and the amounts of class-time (in minutes) assigned to a range of 12 classroom activities, which included practical activities such as demonstrations, class experiments, individualised experiments, field trips, etc. and non-practical activities such as lecturing, question and answer sessions, discussions, etc.

RESULTS

For each observer, an inter-observer agreement index, R_a (Eggleston *et al.*, 1975), was calculated for each of the 21 categories of the SLAS instrument. For each category an intra-observer reliability index was also calculated using Medley and Mitzel's (1958) analysis of variance design. The results are shown in Table 1.

For each classroom visited, a profile of the probabilities of observing of the 21 SLAS categories was computed across the six representative lessons. A hierarchical, centroid cluster-analysis, with the use of a distance coefficient as a measure of similarity between profiles, was then employed as a data-reduction procedure to group together classrooms sharing similar profiles with respect to the 21 SLAS categories. Recent research (Eggleston *et al.*, 1976; Aitkin *et al.*, 1981) has established that a degree of overlap between different groups can be expected when cluster-analysis techniques are applied to profiles of classroom events. For this reason, a computer program was developed which would allow for iterative relocation of classrooms during the clustering process, and thus allow correction of a poor initial partition (Hacker and Lyne, 1982).

TABLE 1

THE OBSERVER RELIABILITY DATA

	R_a			
Category	Observer 1	Observer 2	Observer 3	R_g
A	1·00	1·00	1·00	1·00
B	0·77	0·83	0·77	0·88
C	0·83	0·77	0·83	0·90
D	0·88	0·83	0·94	0·86
E	0·88	0·83	0·83	0·92
F	0·72	0·72	0·83	0·71
G	1·00	1·00	1·00	0·96
H	1·00	0·94	0·94	0·90
I	1·00	1·00	0·94	0·85
1	0·77	0·94	0·88	0·71
2	0·88	1·00	0·94	0·73
3	0·60	1·00	0·88	0·67
4	0·71	0·60	0·77	0·92
5	0·88	1·00	0·88	0·95
6	0·71	0·77	0·77	0·88
7	0·83	1·00	0·83	0·86
8	0·83	0·83	0·94	0·90
9	0·77	0·54	0·83	0·84
10	0·88	0·94	0·88	0·96
11	1·00	0·94	1·00	0·95
12	1·00	0·94	1·00	0·72

KEY: 1. Categories A to I and 1 to 12 are defined in Figure 1.
2. R_a is the inter-observer reliability measure.
3. R_g is the intra-observer reliability measure.

Using procedures outlined by Everett (1974) an optimal solution of six groups was derived by inspection of the dendrogram resulting from the clustering process. The characteristics of these six groups, with respect to probabilities of occurrence of the 21 SLAS categories, are provided in Figure 2. Figure 3 shows the distributions of classrooms at each of five levels of schooling (lower, middle and upper primary; lower and upper secondary) into these six groups.

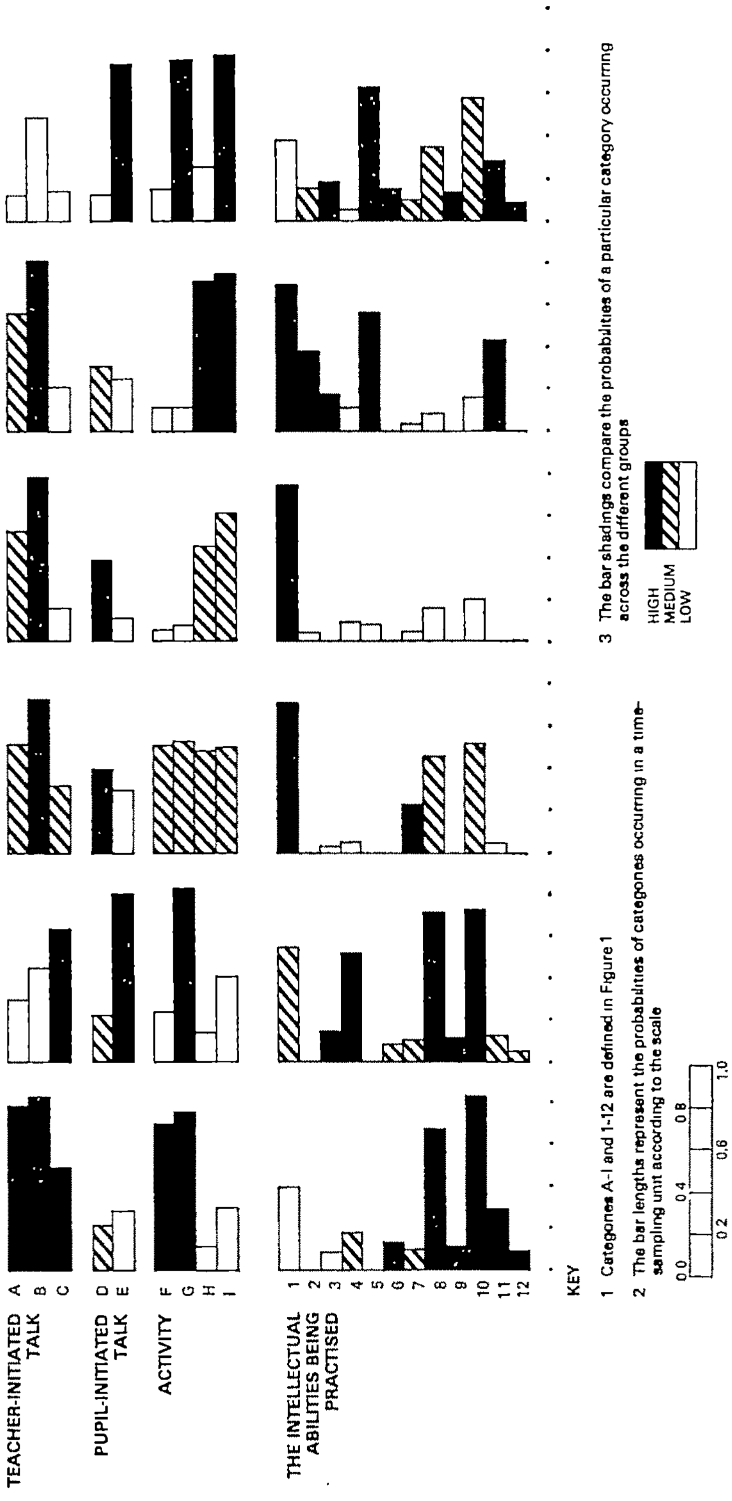
DISCUSSION AND CONCLUSIONS

Figures 2 and 3 and the lesson abstracts provided an empirical basis for a descriptive account of the intellectual development of children in the classrooms studied. A three-tier model is postulated and this model is summarised in Table 2.

Groups 1 and 2 comprise the first tier of this hierarchy. The intellectual focus for classrooms comprising these groups is the development of basic syntactic skills and application of these skills to concrete problem-solving activities. Group 3 and 4 classrooms constitute the middle tier of the hierarchy. Here emphasis is placed on the development and consolidation of a broad scientific informational base as a prelude to higher-order substantive development. Groups 5 and 6 comprise the third tier of the hierarchy. These classrooms are concerned with the development of theoretical constructs of increasing abstraction and power, their data-summarising capacities, and their hypothesis-generating characteristics.

FIGURE 2

THE CHARACTERISTICS OF THE GROUPS WITH RESPECT TO THE PROBABILITIES OF THE SLAS CATEGORIES OCCURRING DURING A TIME-SAMPLING UNIT



KEY

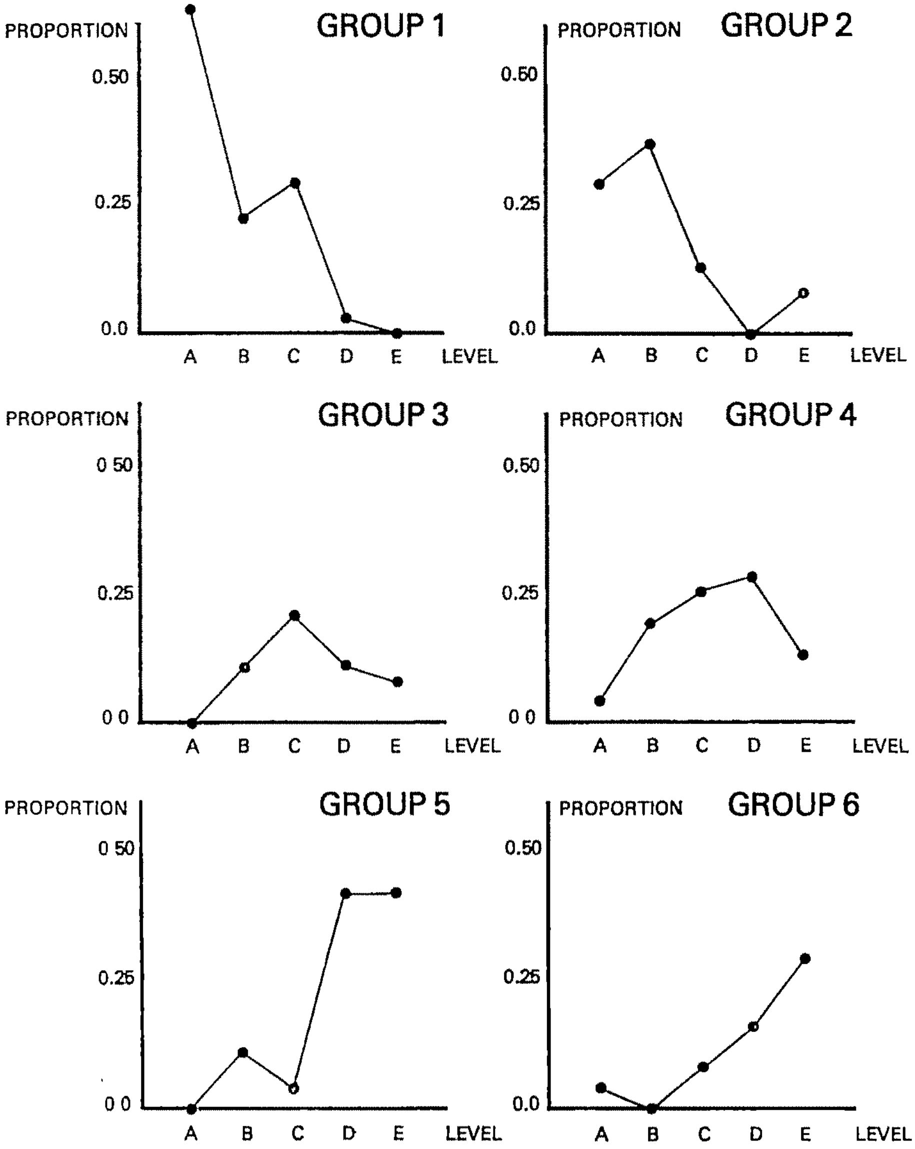
1 Categories A-I and 1-12 are defined in Figure 1

2 The bar lengths represent the probabilities of categories occurring in a time-sampling unit according to the scale

3 The bar shadings compare the probabilities of a particular category occurring across the different groups

FIGURE 3

PROPORTIONS OF THE CLASSROOMS AT EACH LEVEL OF SCHOOLING COMPRISING EACH GROUP



KEY TO THE LEVELS OF SCHOOLING

A Lower Primary (Years 1 and 2)
B Middle Primary (Years 3 to 5)
C Upper Primary (Years 6 and 7)
D Lower Secondary (Years 8 to 10)
E Upper Secondary (Years 11 and 12)

TABLE 2

A HIERARCHY OF INTELLECTUAL DEVELOPMENT IN SCIENCE

Group	Schooling level of modal proportion	The intellectual abilities emphasised in the classroom
1	Lower-primary	The basic syntactic skills; making measurements and observations, interpreting observations and combining these interpretations in an inferential manner. Practising elementary experimental procedures, formulating simple hypotheses about scientific phenomena and devising univariate experiments to test these hypotheses.
2	Middle-primary	Consolidation of syntactic abilities and application of these skills to practical problem-solving activities. Increasing emphasis on the informational aspects of science.
3	Upper-primary	Extension of the field of application of syntactic skills to verifying and confirming scientific information. Considerable emphasis on the assimilation of factual information; acquiring, recalling and confirming facts; identifying and describing scientific substances, materials and equipment.
4	Lower-secondary	Consolidation and extension of a scientific informational base as a prelude to higher-order substantive development. The genesis of concept development.
5	Lower-secondary/ Upper-secondary	Development of progressively more abstract theoretical constructs. Emphasis on the data summarising characteristics of these constructs and their application to problem-solving activities. Interpreting recorded observational data.
6	Upper-secondary	Emphasis on the hypothesis-generating characteristics of theoretical constructs. Formulating hypotheses about significant scientific problems and designing controlled, multivariate, empirical tests. Continuing emphasis on the powerful data-summarising capacities of higher-order theoretical constructs and their applications to problem-solving activities.

The hypothesis that the distribution of classrooms across these three tiers varied with the levels of schooling (see Figure 3) was tested via a contingency table analysis (Nie *et al.*, 1975). A chi-squared result of 81·9 on 8 degrees of freedom supported this hypothesis ($P < 0{\cdot}0001$). Tier I classrooms (groups 1 and 2) were mainly found at the lower and middle primary levels. Tier II classrooms (groups 3 and 4) were found mainly at the upper primary and lower secondary levels, whereas Tier III classrooms (groups 5 and 6) were generally found in secondary schools.

For the abilities studied, the results are generally consonant with the notion of gradual, quantitative changes in pupil thinking, rather than with the notion of discrete, qualitative stages with emergent intellectual functions. Pupils as young as 5 or 6 years of age, at the lower-primary level, were observed engaging in intellectual behaviours classified under no fewer than 10 of the 12 generic categories of pupil behaviours studied. They were observed to make hypotheses about scientific phenomena and to devise hypothetical, univariate empirical tests for their ideas. These findings are at variance with the predictions of Piagetian theory in its original form, where the ability to develop hypotheses is a characteristic of formal operational thinking (Inhelder and Piaget, 1958). Also, it was clear from lesson abstracts that the hypothetical thinking observed in the classrooms visited went beyond the ' theories-in-action ' which have been proposed by Karmiloff-Smith and Inhelder (1975) as a revision to the original Genevian position.

There is further evidence of dissonance between the results of this study and the predictions of Piagetian theory. Children at lower-primary school level were frequently observed to make observations, to derive propositions by interpreting these observations and then to combine these propositions in an inferential manner. These findings are supportive of Bryant's (1977) assertion that Piaget's view of children and inferences is far too pessimistic and that young children can and do engage in inferential thought processes.

Of the 12 generic categories of scientific thought studied, perhaps the incidences of categories 2 and 5 might be considered as being consistent with the notions of emergent, intellectual behaviours. In the classrooms observed, few serious efforts were made to develop and apply concepts, principles and models prior to the lower-secondary and upper-secondary levels of schooling. These two categories subsume many of the 12 logical skills which Piaget regards as providing a sufficient description of advanced intellectual development, and, to this extent, the results are consonant with Piagetian theory.

One substantial advantage in studying cognitive growth via natural behaviour sequences in the classroom is that it is possible to explore relationships between the profiles of intellectual abilities stressed in the classroom and the preferred instructional strategies of teachers. An analysis of the forms of the interactions across the six groups (see Figure 2, categories A to I) and an analysis of lesson abstracts revealed some interesting and unexpectedly complex variations in instructional strategies as cognitive development progressed. In group 1 classrooms, where the development of basic syntactic skills is stressed, teacher-directed inquiry strategies were popular. There were high probabilities of teacher questions, statements and directives occurring in the classroom and frequent interactions of both teacher and pupil with science resource materials. Class experiments, question and answer sessions and demonstrations were the most popular classroom activities (51 per cent, 16 per cent and 12 per cent of class time, respectively) and the teacher mediated frequently between pupils and materials in the learning process; all learning experiences were closely managed and supervised.

In group 2 classrooms, where the focus shifts to the application of syntactic skills to practical, concrete problem-solving, there was a move from teacher-directed towards pupil-directed inquiry strategies. Pupils worked more independently of the teacher and discussed their learning experiences amongst themselves to a much greater extent. Class experiments, individualised experiments and project work were the most popular classroom activities (48 per cent, 15 per cent and 14 per cent of class time, respectively).

Increasing emphasis on the acquisition of scientific information at Tier II of the hierarchy was accompanied by a shift in emphasis from inquiry to expository instructional strategies. In group 3 classrooms, teacher statements became the most probable types of interactions to occur in these classrooms and there was increased emphasis on the use of multi-media materials. Lecturing and question and answer sessions accounted for roughly 50 per cent of class time, whereas demonstrations and class experiments were popular practical activities (20 per cent and 18 per cent of class time, respectively). However, the experiment was generally employed in its illustrative role, to confirm or verify information, rather than as a vehicle for inquiry, this being confirmed by the absence of speculative or inferential categories of pupil behaviour in the classrooms comprising this group.

In group 4 classrooms, the main intellectual emphasis is on development of a scientific informational base. Here expository, non-practical teaching strategies were also dominant: lecturing and question and answer sessions accounted for 66 per cent of class time.

Higher-order substantive development is the main focus of Tier III classrooms. In group 5 classrooms, the development of concepts, principles and models and their applications to problem-solving activities was achieved mainly by expository strategies. The high probabilities of teacher statements and both teacher and pupil interactions with multi-media materials reflected the popularities of both lecturing and theoretical problem-solving sessions (43 per cent and 35 per cent of class time respectively).

In group 6 classrooms, as more attention is given to the hypothesis-generating characteristics of theoretical constructs, the popularity of pupil-directed inquiry approaches increased. Classrooms were characterised by low probabilities for teacher-directed categories of behaviour and higher probabilities for pupil-directed categories. Pupils worked more independently of the teacher than at any of the preceding levels. Practical teaching strategies were generally employed to develop hypothetico-deductive reasoning skills and facility in designing experiments; class and individualised experiments accounted for 45 per cent and 27 per cent of class time, respectively. Experiments designed at this level were characterised by greater awareness of the influences of more than one variable on the outcome of the experiment and more effective control of extraneous variables.

Overall, inquiry strategies were more popular for developing syntactic skills in the primary school science lessons observed, whereas the instructional strategies adopted at the secondary level were generally more consistent with the expository tactics recommended by Ausubel (1968) and Gagné (1977). The internalisation of new facts, concepts, principles and models was more frequently achieved by direct instruction rather than by the cognitive restructuring process favoured by Piaget (1964) and Bruner (1966). Perhaps these instructional preferences go some way towards explaining, for the population studied, very low adoption rates of secondary science curricula based extensively on cognitive interactionist learning strategies.

Fifty-seven of the 60 secondary school classrooms studied were found in groups comprising Tiers II and III of the hierarchy. These classrooms (years 8 to 12) were classified according to whether a biological science topic or a physical science (physics and chemistry) topic was being studied. The hypothesis that the distributions of classrooms into these two tiers was dependent on whether a biological or physical science topic was being studied was tested. A chi-squared value of 24·8 on 1 degree of freedom supported this hypothesis ($P < 0{\cdot}0001$).

Classrooms where chemistry or physics topics were being studied predominated at the upper tier of the hierarchy (83 per cent and 85 per cent of the secondary classrooms in groups 5 and 6, respectively), whereas more classes studying biology topics were found in groups comprising the middle tier of the hierarchy. In a study of science classroom interactions at the secondary school level, Eggleston (1980) has also reported lower levels of intellectual demand on biology students than on pupils of the physical sciences. It seems reasonable to conclude that both sets of results reflect differences in the states of development of these science disciplines as they are presented at the secondary school level.

The hierarchical model of intellectual development which has resulted from this study maps the cognitive growth of science pupils throughout the primary and secondary school. The model is explicit about the types of intellectual behaviours in which pupils engage and, for this reason, it provides a testable account of intellectual progress. Its applicability to other populations of science classrooms could be tested. The profiles of intellectual behaviours reported are descriptions of the steps which practising science teachers see as the optimal path between children's initial experiences of science and achievement of the intellectual behaviours of a scientist, given the practical context of the school. They should not be interpreted as setting upper limits on pupils' capacities. Indeed, Case's (1974) work suggests that children of 7 to 8 years of age can acquire control of variables schemes, whereas these skills were not formally developed in the population studied until the latter years of secondary schooling.

As well as being testable, the proposed model of intellectual development in science classrooms is useful to both theorists and practitioners in education. It provides cognitive theorists with an empirical base for comparison with theoretical prescriptions; it provides teachers and curriculum managers with information about the kinds of intellectual development which can be expected to occur in science classrooms at different levels of schooling; and it provides curriculum developers with information which is useful in selecting and sequencing learning experiences and courses of study.

Information about the preferred instructional strategies of the science teaching population is also useful. There is a growing body of evidence to suggest that high levels of dissonance between the strategies recommended by curriculum projects and the preferences of classroom teachers lead to rejection or ' adaption ' of curricula at the classroom level. It is only by gleaning information about classroom practices that this dissonance level can be monitored and controlled.

ACKNOWLEDGMENT.—This work was supported by an operating grant from the Education Research and Development Committee, The Commonwealth Department of Education, Canberra, ACT, Australia.

REFERENCES

AITKIN, M., BENNETT, S. N., and HESKETH, J. (1981). Teaching styles and pupil progress: a re-analysis *Br. J. educ. Psychol.*, 51, 170-186.

AUSUBEL, D. P. (1968). *Educational Psychology*. New York: Holt, Rinehart and Winston.

BRAINERD, C. J. (1978). *Piaget's Theory of Intelligence*. Englewood Cliffs, N.J.: Prentice-Hall.

BROWN, R. G., and DESFORGES, C. (1979). *Piaget's Theory: A Psychological Critique*. London: Routledge and Kegan Paul.

BRUNER, J. S. (1960). *The Process of Education*. Cambridge, Mass.: Harvard University Press.

BRUNER, J. S. (1966). *Towards a Theory of Instruction*. Cambridge, Mass.: Harvard University Press.

BRYANT, P. E. (1977). Logical inferences and development. In GEVER, B. A. (Ed.), *Piaget and Knowing*. London: Routledge and Kegan Paul.

CASE, R. (1974). Structures and strictures: some functional limitations on the course of cognitive growth. *Cogn. Psychol.*, 6, 544-573.

CASE, R. (1975). Gearing the demands of instruction to the developmental capacities of the learner. *Rev. educ. Res.*, 45, 59-87.

DEWEY, J. (1902). The child and the curriculum. In ARCHAMBAULT, R. D. (Ed.), *John Dewey on Education*. New York: Random House.

DRIVER, R. (1982). Children's learning in science. *Educ. Analysis*, 4, 69-79.

EGGLESTON, J. F. (1980). Studies of science teaching processes. In MCFADDEN, C. P. (Ed.), *World Trends in Science Education*. Halifax, N.S.: Atlantic Institute of Education.

EGGLESTON, J. F., and GALTON, M. J. (1981). Reply to the article by Dunkerton and Guy. *Eur. J. sci. Educ.*, 3, 316.

EGGLESTON, J. F., GALTON, M. J., and JONES, M. E. (1975). *A Science Teaching Observation Schedule*. London: Macmillan Education.

EGGLESTON, J. F., GALTON, M. J., and JONES, M. E. (1976). *Processes and Products of Science Teaching*. London: Macmillan Education.

EVERETT, B. (1974). *Cluster Analysis*. London: Heinemann Educational.

FENSHAM, P. J., GARRARD, J., and WEST, L. H. T. (1981). The use of cognitive mapping in teaching and learning strategies. *Res. sci. Educ.*, 11, 121-129.

FLAVELL, J. H. (1963). *The Developmental Psychology of Jean Piaget*. London: Van Nostrand.

GAGNÉ, R. M. (1977). *The Conditions of Learning*. (3rd edit.). New York: Holt, Rinehart and Winston.

HACKER, R. G. (1980). Cognitive interactions in science: classroom practices and some prescriptions of theories of learning. *Alberta J. educ. Res.*, 26, 25-35.

HACKER, R. G. (1982). *The Science Lesson Analysis System*. Research report, Department of Education, The University of Western Australia.

HACKER, R. G., HAWKES, R. L., and HEFFERNAN, M. K. (1979). A cross-cultural study of science classroom interactions. *Br. J. educ. Psychol.*, 49, 51-58.

HACKER, R. G., and LYNE, A. (1982). *Shufter: A FORTRAN IV programme for Agglomerative Hierarchical Cluster Analysis with Iterative Relocation of Entities.* Research report, Department of Education, The University of Western Australia.

INHELDER, B., and PIAGET, J. (1958). *The Growth of Logical Thinking from Childhood to Adolescence.* New York: Basic Books.

KARMILOFF-SMITH, A., and INHELDER, B. (1975). If you want to get ahead, get a theory. *Cognition*, 3, 195-212.

KLAHR, D., and WALLACE, J. G. (1976). *Cognitive Development: An Information Processing View.* New York: Lawrence Erlbaum.

KLOPFER, L. E. (1971). Evaluation of learning in science. In BLOOM, B. S., HASTINGS, J. T., and MADAUS, G. F. (Eds.), *A Handbook on Formative and Summative Evaluation of Student Learning.* New York: McGraw-Hill.

LAWSON, A. E., KARPLUS, R., and ADI, H. (1978). The acquisition of propositional logic and formal operational schema during the secondary school years. *J. Res. sci. Teaching*, 15, (6), 465-478.

LOVELL, K. (1980). The relevance of cognitive psychology to science and mathematics education. In ARCHENHOLD, W. F. *et al.* (Eds.), *Cognitive Development Research in Science and Mathematics.* Leeds: University of Leeds Press.

MEDLEY, D. M., and MITZEL, H. E. (1958). Application of analysis of variance to the estimation of the reliability of observation of teachers' classroom behaviour. *J. exper. Educ.*, 27, 23-25.

NIE, N. H., HULL, C. H., JENKINS, J. G., STEINBRENNER, K., and BENT, D. H. (1975). *Statistical Package For The Social Sciences.* New York: McGraw-Hill.

OSBORNE, R. J., and BELL, B. F. (1983). Science teaching and children's views of the world. *Eur. J. sci. Educ.*, 5, 1-14.

PIAGET, J. (1964). Development and learning. *J. Res. sci. Teaching*, 2, 176-186.

PIAGET, J. (1971). *Biology and Knowledge.* Chicago: University of Chicago Press.

SCHWAB, J. J. (1960). What do scientists do? *Behavioral Science*, 5, 1-27.

SCHWAB, J. J. (1962). The concept of the structure of a discipline. *Educ. Record*, 43, 197-205.

SCHAB, J. J., and BRANDWEIN, P. F. (1964). *The Teaching of Science.* Cambridge, Mass.: Harvard University Press.

SHAYER, M. (1980). The match of science curriculum to the learner in the middle and secondary school. In ARCHENHOLD, W. F. *et al.* (Eds.), *Cognitive Development Research in Science and Mathematics.* Leeds: University of Leeds Press.

SHAYER, M., and ADEY, P. (1981). *Towards a Science of Science Teaching.* London: Heinemann Educational.

TRABASSO, T. (1977). The role of memory as a system in making transitive inferences. In KAIL, R. V., and HAGEN, J. W. (Eds.), *Perspectives on the Development of Memory and Cognition.* Hillsdale, N.J.: Lawrence Erlbaum.

(*Manuscript received* 14*th September*, 1983)

Br. J. educ. Psychol., 54, 152-159, 1984

THE DEVELOPMENT OF EXPRESSIVITY IN DRAWING

By S. W. IVES

(*Spectrum Training Corporation and Harvard Project Zero*)

SUMMARY. To trace the development of expressivity in children's drawings, 128 children, 16 each at the mean ages of 4, 5, 7, 9, 11, 13, 16, 20 were asked to create a series of drawings depicting expressive qualities (e.g., happy, sad) in both trees and lines. Drawings were first scored in terms of the accuracy of both expression and representation (trees, lines). Successful expressive drawings were then scored in terms of strategy utilised: literal representation (e.g., crying face for sad), abstract expression (e.g., drooping lines for sad), and content expression (e.g., a barren tree for sad). Finally, drawings scored as abstract expression were rated as to the types of abstract contrasts used (e.g., angle/curve, big/little, upward/downward). The results document a developmental trend from an initial ability to produce abstract expression based, in part, on amodal similarities (e.g., big for loud) toward the production of more complex multi-dimensional types of abstract expression, and increasing integration of representational ability and expressive ability in the form of content expression.

INTRODUCTION

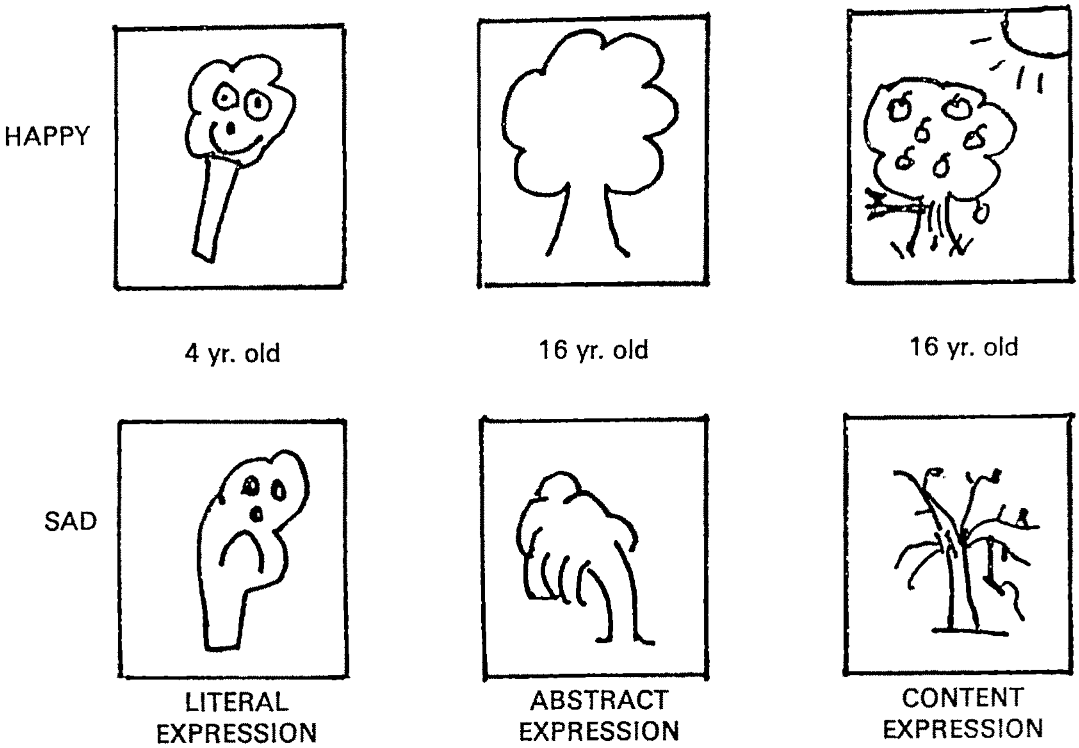
WHILE an increasing number of researchers have begun to study the cognitive aspects of drawing (e.g., Goodnow, 1977; Willats, 1977; Freeman, 1980; Ives and Houseworth, 1980), the general framework for these investigations has been that of literal representation. This is a limited view, however, as a drawing is much more than what it represents and, in fact, it need not be directly representational at all. As Goodman (1968) argues, a drawing can literally represent an object or scene (e.g., a seascape) but it can also express both psychological moods (e.g., sadness) and non-visual sensory qualities (e.g., quietness). Such qualities can be portrayed in a number of ways. They can be represented literally (e.g., a crying face) or non-literally through metaphorical expression. There are two basic ways that pictures can metaphorically express qualities such as moods: through *abstract* (non-representational) features (e.g., dark colours or drooping lines can express sadness, large shapes can express loudness) and through *content* (e.g., a barren tree or a broken swing can express sadness). In contrast to ordinary representation (e.g., a drawing of a sad person represents sadness), both abstract and content expression depict expressive qualities metaphorically (see Figure 1).

Expressive features are frequently used in drawings. Werner and Kaplan (1963) have shown that adults both generally agree on their interpretation and are consistent in their use (e.g., roundness and upwardness for positive qualities and angularity and downwardness for negative qualities). Whether this agreement is universal or culture-specific awaits further study. However, while we know an increasing amount about the development of representation in drawing, little is currently known about expression. In one study, Carothers and Gardner (1979) found that children did not acquire expressive ability until age 10 or 12. However, they only investigated expression through content (e.g., a stormy day to express sadness, a sunny day to express happiness). While Scarlett *et al.* (1980) found some ability to depict expressive qualities by age 5, the children in their study only used literal representation (e.g., sadness depicted through tears) and not abstract or content expression.

A comprehension study by Gardner (1974) indicates that 4-year-olds can correctly distinguish the qualities expressed by drawings using abstract expression (e.g., curved lines for " happy," jagged lines for " sad "). Indeed, there may even be a predisposition for this expressive sensitivity in infancy. Wagner *et al.* (1981) found an ability to make " metaphorical " matches which are not based on physical or spatial-temporal covariance (e.g., an ascending line and an ascending auditory tone) at age 11

FIGURE 1

THE THREE STRATEGIES SCORED IN EXPRESSIVE DRAWINGS



months. The above results suggest that the distinction between abstract and content expression, overlooked in the Carothers and Gardner (1979) study, may be a crucial one.

The present study of the development of expressivity in drawing is designed to address a number of previously unexplored questions. First, how does development in the ability to express both moods (e.g., sadness) and non-visual qualities (e.g., quietness) proceed? Second, what is the effect of depicting expressive qualities through both representational (e.g., sad trees) and non-representational (e.g., sad lines) drawings? Third, how does expressive ability relate to the ability to use drawings for representational (i.e., trees) purposes? Fourth, how are the three strategies of literal representation, abstract expression, and content expression used to depict expressive qualities? Finally, what is the consistency in which children use specific types of abstract expression (e.g., round *versus* angular lines, upward *versus* downward lines)?

METHOD

Sample

A total of 128 subjects, eight males and eight females, in each of eight age groups (approximate mean ages: 4, 5, 7, 9, 11, 13, 16, and 20), participated in the study. They were randomly selected from schools serving middle class populations. The 4-year-olds were tested individually, all 20-year-olds were tested in one group and all other age groups were tested in groups of four. In group situations, subjects were seated so as to prevent copying. The testing was carried out on two separate days for all ages except the 20-year-olds who completed all tasks in one session.

Procedure

Children were asked to sketch the following 12 drawings. To investigate the ability to express psychological moods, subjects were asked to draw a happy tree, a sad tree, and an angry tree. To investigate the ability to express non-visual sensory qualities, they were asked to draw a quiet tree, a loud tree, and a hard tree. To determine whether expressiveness is easier to achieve in a non-representational picture, subjects were asked to draw a happy or sad (etc.) " line." These drawings were then scored on three separate levels of analysis to answer the questions raised in the introduction (see scoring).

Each session began with a warm-up exercise in which subjects were given an 8 in. × 7 in. sheet of paper and told to draw anything they wished with a black felt tip pen. Following this, subjects drew a total of two sets of drawings each day, each of which consisted of three separate drawings. For example, one group was first directed to draw a happy, a sad, and an angry tree " so someone else could tell them apart." Then, they drew quiet, loud, and hard lines. On the second day they drew two more sets of three drawings. Drawings were done on 8 in. × 21 in. sheets of paper which were divided into three 8 in. × 7 in. sections. The order in which the four sets of drawings were presented was counterbalanced within each age group.

Scoring

The drawings of each subject were presented to one of two pairs of judges in paired comparisons: happy-sad, sad-angry, quiet-loud, loud-hard (trees and lines) for a total of eight pairs of drawings per child. The judges were blind to the subject's age and sex. Since judgments were made between, for example, happy-sad and sad-angry, the same judge could not score both of these contrasts by the same subject since the sad one would be remembered as the second judgment was made. Therefore, each pair of judges scored half of all the data within each category, avoiding scoring the same drawing by the same subject twice. The judges were first asked to determine both representational content (i.e., tree *vs.* lines) (91 per cent reliability) and expressive content (e.g., sad *vs.* angry) (84 per cent reliability). To eliminate forced guesses and a high chance level, judges could also rate the pairs as " no difference." The drawings which were scored correctly for expressive qualities were then scored in a separate and subsequent analysis for the strategies used to express the mood or sensory quality: either literal representation, abstract expression, content expression or some combination of these three (89 per cent reliability). For those pairs which the judges had been unable to distinguish the expressive quality, the reason for failure was scored as either " no difference " or " inappropriate attempt " (89 per cent reliability).

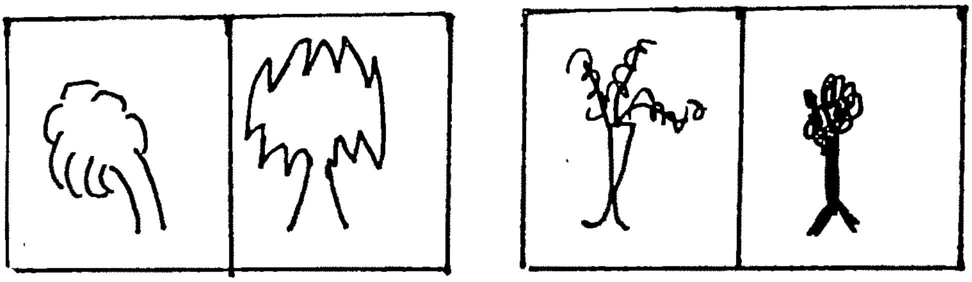
The drawings judged correctly in terms of expressive qualities which used abstract expression were scored a third time to determine the type or types of abstract expression used. Again, this analysis was separate and subsequent to the two previous analyses. Based on previous work and a preliminary analysis of the data, five categories were used in scoring: (1) angular *vs.* curved, (2) upward *vs.* downward, (3) thick *vs.* thin, (4) big *vs.* little, and (5) the page placement of lines or objects (up or down) (see Figure 2). The drawings were scored by judges blind to age and sex using these five categories as well as an " other " category (87 per cent reliability).

Analysis

The various dependent measures (e.g., accuracy, strategy) were analysed using 8 (Age) × 2 (Sex) × 2 (Subject: tree/line) × 4 (Expressive Contrast: happy/sad, etc.)

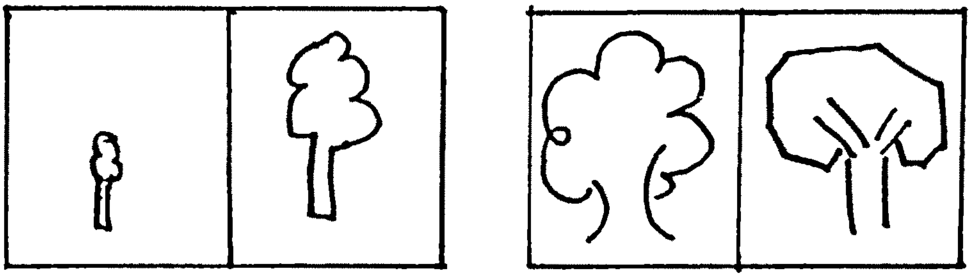
FIGURE 2

THE FOUR PREDOMINANT TYPES OF ABSTRACT EXPRESSION



Sad DOWNWARD — Angry UPWARD
16 yr. old

Loud THIN — Hard THICK
4 yr. old



Quiet LITTLE — Loud BIG
4 yr. old

Loud CURVE — Hard ANGLE
16 yr. old

repeated measures analyses of variance with subject and expressive contrast as the within-subject variables. It should be noted that while these ANOVAs were performed on dichotomous dependent measures (e.g., whether or not a drawing was correct) Lunney (1970) has demonstrated that ANOVAs can be used for such situations provided certain conditions (e.g., degrees of freedom of the error term are greater than 40) are present, and that is the case in this study.

RESULTS

Table 1 presents the results for overall accuracy for expressive qualities at each age level as well as the types of errors and the types of expressive strategies present in drawings correctly interpreted by the judges. Children's ability to depict expressive qualities generally increases with age, $F(7,120) = 13{\cdot}29$, $P < 0{\cdot}001$, with significant differences ($P < 0{\cdot}05$, Scheffé method) between each age level except ages 16 and 20 (see Table 1). However, there is a significant dip in correct scores from age 7 to 9 (70·3 per cent to 59·4 per cent, $P < 0{\cdot}05$, Scheffé method). Even 11-year-olds are not up to the 7-year-old level, while the 13-year-olds achieve an 81·3 per cent accuracy rate. This dip in scores can be accounted for, in part, by an increase in inappropriate attempts at age 9, while the presence of pairs exhibiting no difference continues its decline (see Table 1). This dip is also reflected in the decline of the use of literal

TABLE 1

MEAN PERCENTAGE OF TOTAL RESPONSES

Correct responses by strategy[a]	Age 4	Age 5	Age 7	Age 9	Age 11	Age 13	Age 16	Age 20
Literal only	13	27	20	07	02	14	06	00
Literal and abstract	04	04	09	06	06	11	12	03
Abstract only	19	21	38	40	52	48	50	65
Content expression	00	00	04	07	09	09	26	19
Total correct	37	52	70	59	69	81	92	87
Error types								
No difference	41	19	03	03	03	02	01	01
Inappropriate attempt	23	30	27	38	28	16	07	13

[a] Literal = literal representation, abstract = abstract expression, content expression = content expression either alone or in combination with the other two strategies.

representation at age 9 (29 per cent of correct responses at age 7 and 12 per cent at age 9), while the use of abstract expression continues to increase (53 per cent of correct responses at age 7 and 67 per cent at age 9). Thus children at age 9 are more likely than 7-year-olds to utilise abstract expression and only rarely use literal representation. However, they are also less likely to produce a recognisable expressive contrast.

There are also accuracy differences between the types of expressive qualities depicted. Subjects are more accurate when depicting psychological moods (72 per cent) than when depicting non-visual sensory qualities (65 per cent), $F(1,120) = 2{\cdot}82$, $P < 0{\cdot}01$. In addition, as might be expected, qualities which require a more subtle distinction (angry *vs.* sad and hard *vs.* loud) are more difficult (64 per cent accurate) than those which are opposite ends of a continuum (happy *vs.* sad and quiet *vs.* loud) (72 per cent accurate), $F(1,120) = 3{\cdot}13$, $P < 0{\cdot}005$.

It should be noted that the teams " accuracy " and " correctness " are used to refer to the children's ability to depict the qualities so " someone else can tell them apart." It would be misleading to imply that an expressive drawing, by itself, can be considered correct or incorrect. Correctness is used only to refer to the judges' ability to make proper distinctions.

Trees vs. Lines. A comparison of representational (trees) drawings with non-representational (lines) drawings reveals that judges are equally able to determine the expressive qualities in question. However, as might be expected, subjects use a more diverse pattern of strategies in representational pictures. ANOVAs reveal significant differences ($P < 0{\cdot}001$) between tree and line drawings on the correct use of all three strategies. In line drawings subjects rely primarily on the exclusive use of abstract expression (79 per cent), while in tree drawings, subjects divide their responses among the exclusive use of literal representation (21 per cent), the exclusive use of abstract expression (43 per cent), combinations of these two (13 per cent), and the use of content expression alone or in combination with the other two strategies (23 per cent).

Representational vs. Expressive Ability. As with expressive ability, children's abilities to depict representational content (trees *vs.* lines), increases with age, $F(4,120) = 8{\cdot}33$, $P < 0{\cdot}001$. However, 4-year-olds are 67·2 per cent accurate in representing trees and only 36·7 per cent accurate in conveying expressive content.

The depiction of trees reaches a ceiling (100 per cent accurate) at age 7, whereas development in expressive ability advances more slowly peaking only at age 16 (92 per cent accurate). Moreover, amongst the 4-year-olds there are more recognisable drawings of trees (37·5 per cent) which fail to depict expressive content (e.g., sadness) in a recognisable way than the reverse. However, a number (10·9 per cent) of drawings done at this age depict recognisable expressive qualities and yet fail to depict recognisable trees (e.g., a drawing judged correctly on happiness or sadness but not on whether it is a tree or a line).

Strategies. As children begin to create successful expressive drawings, a complex pattern of strategy use emerges. ANOVAs reveal significant age effects ($P < 0{\cdot}001$) for the use of all three strategies (see Table 1). For 4-year-olds, the use of abstract expression is greater than the use of literal representation (51 per cent *vs.* 34 per cent direct responses). This pattern reverses itself at age 5 (52 per cent of correct literal and 41 per cent of correct abstract). Scheffé post-hoc comparisons reveal significant differences between the use of both strategies by both 4- and 5-year-olds. The use of abstract expression does not continue to decline however, but rather increases in subsequent ages while the use of literal representation declines. The use of content expression first appears at age 7, but only comes into substantial use at age 16.

TABLE 2

TYPES OF ABSTRACT EXPRESSION BY AGE EXPRESSED AS PERCENTAGES

Type of abstract expression[a]	Age 4 (28)[b]	Age 5 (30)	Age 7 (63)	Age 9 (59)	Age 11 (83)	Age 13 (81)	Age 16 (105)	Age 20 (106)
Angle/curve	25[c]	40	43	61	51	41	59	60
Upward/downward	11	40	40	32	40	40	41	41
Thick/thin	25	07	10	12	17	26	33	31
Big/little	54	30	35	25	30	37	32	35
Page placement (up/down)	04	00	00	02	08	04	10	13
Multiple use[d]	14	20	27	32	42	44	57	56

[a] The five types listed account for 97 per cent of the data.

[b] The number of drawings correctly using abstract expression in age group.

[c] Figures are percentages of the correct drawings using abstract expression in each age level which used the type listed (e.g., angle/curve). Note that a drawing can use more than one type.

[d] Percentage of correct drawings using abstract expression in each category which used more than one type (e.g., angle/curve and big/little).

Types of Abstract Expression. Four types of abstract expression (angle/curve; upward/downward; thick/thin; and big/little) account for 91 per cent of the correct instances of abstract expression (see Figure 2 for examples) and page placement accounts for an additional 6 per cent. These four principal types are found in drawings by all the age groups although there are some developmental differences in the frequency of their use (see Table 2). Young children tend to use only one type of abstract expression within a drawing and there is a continuous increase with age in the use of combinations of types of abstract expression within the same drawing (e.g., combination of angle/curve and big/little) (see Table 2).

DISCUSSION

This study indicates that children as young as 4 years can depict expressive qualities in a meaningful way using abstract forms. Their use of such abstract linear forms as angularity and largeness can be correctly interpreted by adults. Moreover, the results suggest that representational ability is not a prerequisite for expressive

P3542

ability. Based on the present findings, development in the use of expressivity in drawing can be tentatively characterised as follows. Expressivity is acquired in the pre-school years where there is an initial ability to produce abstract expression based, in part, on what appears to be amodal or cross-modal similarities (e.g., big for loud, small for quiet). In these early drawings, children generally use only one dimension (e.g., size difference only) to abstractly convey expressivity. As children get older they use more complex, multi-dimensional types of abstract expression as well as content expression.

The strategy of content appears more complex than abstract expression as it combines the two elements of representation and expression. In content expression, the drawing cannot simply be based on a cross-modal or amodal match between the expressive quality and a linear configuration. Instead, the drawer must conceive a situation which metaphorically captures the expressive quality (e.g., a broken swing for sadness) and then create a representation of that situation, a more complex process requiring a more conscious use of the medium (see Figure 1). The relative difficulty of producing abstract and content expression parallels the findings in drawing comprehension tasks. Gardner (1974) found that 4-year-olds had some ability to understand cross-modality matches based on abstract expression, whereas Kogan *et al.* (1980) found that the ability to make matches on what can be called content expression appears later, developing throughout the elementary school years.

In the present drawing tasks, four major types of abstract expression are used consistently at all ages. This surprising commonality across ages is also consistent with Werner and Kaplan's (1963) extensive work with adults' line drawings. As mentioned, these authors found, for example, that normal adults commonly used upward for positive and downward for negative, curved for positive and angular for negative. These contrasts generally appear to reflect cross-modal comparisons, e.g., big (visual) for loud (auditory). These comparisons could result from conscious cross-modal links or perhaps from the presence of an internal amodal information processor. In this latter case, the more general quality of increased intensity could be expressed either visually (big) or auditorily (loud). Wagner *et al.* (1981) have argued for the latter explanation in interpreting their findings of infants' ability to make " metaphorical " matches. Bower (1974) and Gibson (1966) have also claimed there is an initial " supermodal " spatial reference system allowing for early auditory-visual co-ordination. The presence of this initial capacity could account for the ability of some 4-year-olds to utilise abstract expression within their drawings. However, the graphic forms used by children in this study could also be derived from a variety of other sources (e.g., original inventions, cultural conventions, school curriculum). The present study was designed to examine children's meaningful use of expressive forms, a relatively unexplored area. Further work will have to be done to determine the origins of these forms.

ACKNOWLEDGMENTS.—This research was supported by a Wheelock College faculty development grant to the author and grants to Harvard Project Zero from the following institutions: Carnegie Corporation, National Institute of Education (G-78-0031), National Science Foundation (BNS 79-24430), and Spencer Foundation. The author would like to thank the staff and students of the Soldier's Field Park Child Care Center, Boston, the Lincoln Nursery School, Lincoln, Mass., the Lynnfield, Mass., public schools, the Weston, Mass., public schools, and Boston College. Appreciation is expressed to Susan Arrington, Ellen Christie, Karen Haper, Noah Heilman, Deborah Mercer for assistance in data collection and Eve Mendelsohn, Martha Vibbert, and Line Ecker for their assistance in scoring as well as their comments. Thanks also go to Howard Gardner, Erin Phelps, George Scarlett, Nancy R. Smith, Doris Sponseller, and Ellen Winner for their helpful suggestions at various stages of this project. Separate portions of this project were presented to the Society for Research in Child Development, Boston, April, 1981, and the Jean Piaget Society, Philadelphia, May, 1981.

REFERENCES

BOWER, T. G. R. (1974). *Development in Infancy*. San Francisco: Freeman.

CAROTHERS, T., and GARDNER, H. (1979). When children's drawings become art: the emergence of aesthetic production and perception. *Dev. Psychol.*, 15, 570-580.

FREEMAN, N. H. (1980). *Strategies of Representation in Young Children*. New York: Academic Press.

GARDNER, H. (1974). Metaphors and modalities: how children project polar adjectives onto diverse domains. *Child Dev.*, 45, 84-91.

GIBSON, J. J. (1966). *The Senses Considered as Perceptual Systems*. Boston: Houghton Mifflin.

GOODMAN, N. (1968). *Languages of Art*. Indianapolis: Bobbs Merrill.

GOODNOW, J. J. (1977). *Children Drawing*. Cambridge, Mass.: Harvard University Press.

IVES, W., and HOUSEWORTH, M. (1980). The role of standard orientations in children's drawings of interpersonal relationships: aspects of graphic feature marking. *Child Dev.*, 51, 591-593.

KOGAN, N., CONNOR, K., GROSS, A., and FAVA, D. (1980). Understanding visual metaphor: developmental and individual differences. *Monog. Soc. Res. Child Dev.*, 45, (Serial No. 183).

LUNNEY, G. H. (1970). Using analysis of variance with a dichotomous dependent variable: an empirical study. *J. educ. Measur.*, 7, 263-269.

SCARLETT, G., FUCIGNA, C., WARE, G., and NELSON, M. (1980). Expressivity in early childhood. Paper presented at the American Educational Research Association, Boston.

WAGNER, S., WINNER, E., CICCHETTI, D., and GARDNER, H. (1981). " Metaphorical " mapping in human infants. *Child Dev.*, 52, 728-731.

WERNER, H., and KAPLAN, B. (1963). *Symbol Formation*. New York: Wiley.

WILLATS, J. (1977). How children learn to draw realistic pictures. *Quart. J. exper. Psychol.*, 29, 367-382.

(*Manuscript received 27th May*, 1983)

Br. J. educ. Psychol., 54, 160-167, 1984

THE RELATIONSHIP BETWEEN MATERNAL RATINGS OF FIRST WORD VOCABULARY AND REYNELL LANGUAGE SCORES

By C. C. CUNNINGHAM AND P. SLOPER
(Hester Adrian Research Centre, University of Manchester)

SUMMARY. Forty-four mothers of children with Down's syndrome (CA mean 61·5 months, range 37-90; MA mean 36·4 months, range 11-63) checked off the words used by their children from a list of 108 ' topic words ' taken from the First Word Language Programme (Gillham, 1979). Significant correlations were found with Reynell Expressive Language Age ($r = 0{\cdot}87$) and Verbal Comprehension Age ($r = 0{\cdot}76$). The importance of formal methods of integrating parents into the assessment process is discussed in relation to the interpretation of sources of variance between parent and professional assessments.

INTRODUCTION

> " Assessment should be seen as a partnership between the teacher, other professionals and the child's parents, in a joint endeavour to discover and understand the nature of his difficulties and needs. . . . Such contacts can only be helped by frankness and openness on all sides."
>
> (DES Circular, 1983)

THE increasing emphasis on parental involvement in early education for children with learning difficulties has included the contribution of parental observations and ratings to assessment procedures. This has been paralleled by a move away from sole reliance on standardised tests to the use of a range of information sources across differing contexts. Thus it is acknowledged that parental observations of their child's behaviour in contexts other than the school is a necessary and invaluable source of data. Newson (1976) stressed this and expressed surprise that " the resource of parents as an information store is so often neglected " (p. 105), adding that since the knowledge possessed by parents is often unstructured and diffuse, professionals need to develop methods which would assist in more easily defining and structuring this knowledge. Parents are also increasing their demands that their contribution to the assessment procedure should be formal and recognised (Family Focus, 1982). In practice, however, there is limited effort by many professionals to do this; parental observations are often dismissed on the grounds that they are biased and likely to overestimate the child's performance. A number of studies have affirmed that mothers tend to rate their pre-school child's developmental level more highly than ratings from diagnostic tests in 5-year-old non-handicapped children (e.g., Hunt and Paraskevopoulos, 1980) and pre-school handicapped children (e.g., Ewert and Green, 1957; Gradel *et al.*, 1981).

This may not be a biased overestimation of the child's ability. Parents frequently complain that their child did not perform at his/her best in the strange setting of the assessment and/or the assessor was unfamiliar. Brown and Semple (1970) reported that behaviour is often minimised when it occurs in unfamiliar situations with a strange tester. Standardised assessment instruments also tend to be inflexible and may not be appropriate for young mentally handicapped children, given their limited repertoire for interaction and their inherent variability (Simeonsson *et al.*, 1980). Typically, the order of estimation of ability appears to be highest with parents (and in our experience, mothers usually higher than fathers) followed by teachers, and lowest in formal diagnostic test administration (Gradel *et al.*, 1981). This order approximates to the amount of contact time with the child and the range of contexts in which to observe the targeted behaviour.

Thus it would seem premature and unproductive to dismiss parental contribution to assessments merely on the generalised assumption of bias. Indeed, whilst studies do find ' overestimation ', they also report high correlations between parental ratings and diagnostic test results (e.g., Gradel *et al.*, 1981). During the course of our longitudinal, home-based early education programme for children with Down's syndrome (Cunningham, 1983) we frequently noted how close mothers' approximations of developmental level were to our assessed levels. Although no measures were taken, we also found that in many cases of disagreements between our assessments and mothers' estimates, re-administration of items or altering the context of the assessment (i.e., allowing mothers to test the child, using familiar objects, or observing the behaviours in less standardised settings) supported the mothers' contention that the child had been seen to do similar behaviours. There were also occasions when this did not happen. These occasions are a potent source of friction between parent and professional and discussion of the possible causes of variance is essential to maintain the parent-professional partnership (Jeffree *et al.*, 1981).

We found that, for many, friction was apparently reduced if parents were given (*a*) information on the nature of standardised normative testing (e.g., the function of standard procedures, how norms are established) and early development, and (*b*) guidance on improving their own observational skills of the child's behaviour in relation to the developmental process. This suggested that formal methods of incorporating parent observations and assessments into the assessment procedure might prove beneficial both for the child and the parent.

We decided to examine this possibility in the area of language assessment for two reasons. Firstly, it is a common source of disagreement between parents and professionals and is a major area of concern for parents. Secondly, early language development is an area which is difficult to assess reliably in the young mentally handicapped child and is particularly reliant upon context.

In this brief study, we decided to compare mothers' ratings of the words used by their young children with Down's syndrome, taken from the First Word Language Programme (Gillham, 1979), with formal assessments of language using the Reynell Developmental Language Scales (RDLS) (Reynell, 1977) carried out in the home.

METHOD

Pilot study

The sample consisted of 18 mothers and their children with Down's syndrome who were involved in an early intervention programme. There were 12 boys and six girls, mean CA 52·7 months (range 44-68), mean MA 26·8 (range 19-46). Using fathers' occupation, seven were rated as social class I+II, four as social class III and seven as social class IV+V.

In order to accustom the child to the experimenter's presence, their mothers were engaged in conversation prior to the testing on the RDLS. During this period the mother was asked if her child used or had used words from a list of 178 taken from ' The First Language Programme ' (Gillham, 1979). Gillham constructed this list from observations on 14 non-handicapped children. Idiosyncratic forms due to articulation problems and alternatives, e.g., bow-wow or doggie for dog, were accepted. Mothers were instructed not to score words only heard as imitations. The mean number of words rated was 115·3 (SD 50·5, range 16-178). The mean Verbal Comprehension Age (VCA), as assessed on RDLS, was 26·9 months (range 18 to 37) and the Expressive Language Age (ELA) was 25·6 months (range 15-37).

Spearman Rank Correlation Coefficients were computed between word frequency and VCA and ELA. The obtained values were r = 0·59 for VCA and r = 0·87 for ELA, both significant at the 1 per cent level. Clearly the correlation with Expressive Language scores is much higher and accounts for 76 per cent of the variance compared to 35 per cent for receptive language. This is to be expected since ' First Words ' used is essentially a measure of expressive language.

Main study

It was decided to extend the study to a larger range of children and to use a more controlled maternal rating procedure. The 108 ' topic words ' (nouns) of the ' First Word Programme ' were typed in alphabetical order on a standard sheet (see Table 1). Prior to the Reynell testing, the parents were given the sheet and asked to tick the words they had heard the child use. Idiosyncratic forms and alternatives were allowed.

TABLE 1

PERCENTAGE FREQUENCY OF WORDS

Percentage	
91-100	ball, bath, car, daddy, mummy, nose, shoe
81-90	baby, bed, book, cat, chair, cup, dog, door, eye, hair, socks, spoon, teeth, toe
71-80	bag, balloon, bike, bird, boat, boy, bus, cake, clock, coat, dinner, drink, ear, fish, fork, hand, hat, key, person's name, pram, stairs, swing, tea, teddy
61-70	apple, banana, bin, bread, brick, bunny, chocolate, cow, dolly, duck, horse, knee, light, lollipop, man, milk, mouth, paper, peg, plane, potty, soap, toast, train, tree
51-60	biscuit, bubble, egg, flower, garden, juice, lorry, moor, pen, scissors, tap, water
41-50	button, cheese, drawer, gate, meat, milkman, mouse, orange, peas, pear, pin, potatoes, pudding, purse, seesaw, spade, van
31-40	bowl, fly, shirt
21-30	blanket, bridge, goose, tractor

Two experimenters collected the data. They established a reliability on the RDLS with six children with Down's syndrome using a tester-observer procedure (Werner and Bayley, 1966). The correlation between the two testers' scores was 0·99. All the testing was carried out in the child's home with mother present. All the children had taken part in the early intervention programme and were being followed up at regular intervals.

Sample

The sample consisted of 44 mothers and their Down's syndrome children, i.e., all the parents of children with a developmental level of 12 months or more and available for testing in May/June, 1980. There were 20 girls and 24 boys. The mean chronological age was 61·5 months (range 37 to 90), mean mental age 36·4 months (range 11 to 63). Using fathers' occupation, 12 were classified as social class I+II, 18 as social class III and 14 as social class IV+V.

RESULTS

Correlations between scores

The mean first word frequency scores (FWS) was 69·8 (range 4 to 108, SD 38·36). The mean VCA was 31·4 months (range 12 to 80, SD 11·72) and the mean ELA was 28·6 months (range 12 to 61, SD 10·68). Because the plots for FWS against VCA and ELA indicated a linear regression, Pearson product-moment correlations were computed (Figure 1). The correlation for FWS and ELA was r = 0·87, $P<0·001$. The correlation for FWS and VCA was r = 0·76, $P<0·001$. Twelve children achieved maximum First Word scores of 107 and 108. These were removed from the data and the correlation for the remaining 32 was FWS and ELA, r = 0·88, $P<0·001$ and for FWS and VCA, r = 0·73, $P<0·001$.

Thus as with the pilot study, the correlation between expressive language age and first words account for the greater percentage of variance (78 per cent). The correlation between VCA and FSW is greater than in the pilot study, accounting for 53 per cent of the variance, but still remains less than that for expressive language level.

Social class

Inspection of Table 2 shows that there is no difference between ELA, VCA and FWS means for the three social class groupings. Correlations between the scores within social class groupings ranged from 0·71 to 0·88.

TABLE 2

DISTRIBUTION OF MEAN ELA, VCA AND FWS FOR SOCIAL CLASS I+II, III AND IV+V

	I+II	III	IV+V
ELA	28·5	29·0	28·3
VCA	32·5	31·5	31·0
FWS	69·8	70·3	69·9

Percentage use of words (Table 1)

The order of frequency of use of words closely corresponded to that reported by Gillham. Because the present sample was older than Gillham's the reported percentage use of many words is higher, and so no direct comparison can be made.

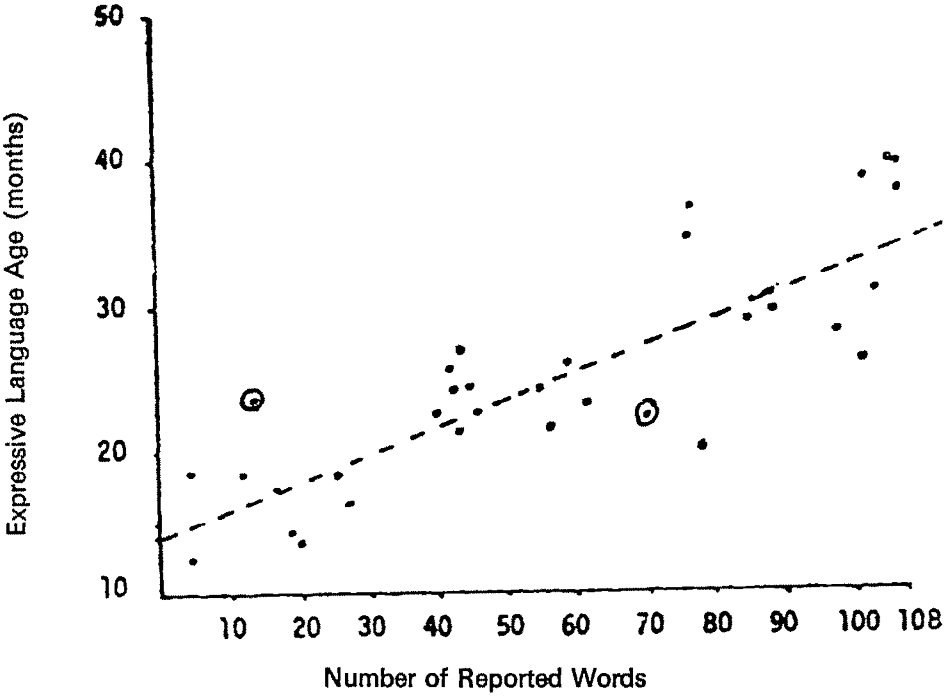
Discrepancies between maternal scores and Reynell assessments

Figures 1 (*a* and *b*) present the scatter of scores and regression line predicting ELA or VCA from number of words reported by mothers, respectively. The 12 children rated as having 107 and 108 words all had high ELA and VCAs in relation to the overall sample and suggested a ceiling effect. Hence they have been excluded. It should also be noted that a score for 20 words of expressive language on the Reynell need not rely on what the experimenter heard but can be augmented by the mother's report. Thus, the two sources are not entirely independent. Further, the Reynell does not just measure vocabulary at this early stage, but also word combinations and grammatical structure. Even with a limited vocabulary many children in the sample did combine words and so achieved relatively higher ELAs than FW scores.

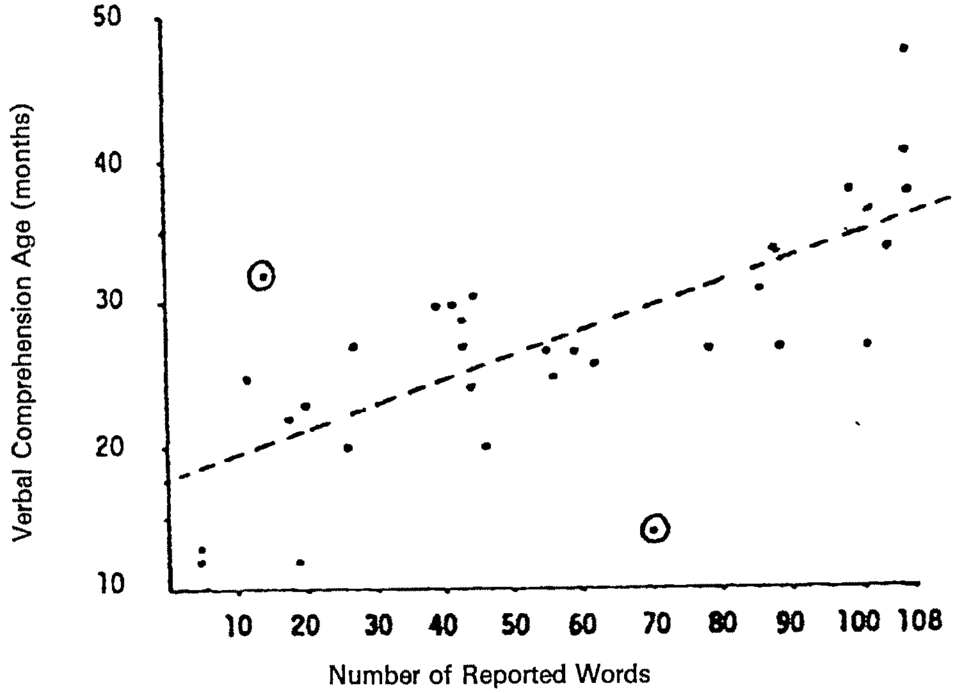
Where ELAs fall below the regression line, this suggests that parent ratings are higher than the assessment and where they fall above the line, lower than the assessment. Inspection of the figures suggests that there is no evidence to support maternal overestimation. Examination of the Reynell scoring sheets revealed that in the case of four children the assessor had noted the session was difficult and the child uncooperative. Three of these cases were of children whose mothers rated the FWSs well below the regression line. This suggests that the discrepancy may be due to the poor reliability of the assessment in that session. Four mothers had been identified by the home visitors of the early education programme, independent of the results, as having

FIGURE 1

(*a*) REGRESSION LINE FOR PREDICTING EXPRESSIVE LANGUAGE AGE FROM NUMBER OF REPORTED WORDS



(*b*) REGRESSION LINE FOR PREDICTING VERBAL COMPREHENSION AGE FROM NUMBER OF REPORTED WORDS



difficulty in establishing a realistic appraisal of the child. Two of these (circled) were found to have the most discrepant results. One, who had been identified as having very negative attitudes, appeared to underestimate the child, and the other, identified as having unrealistically high expectations, to overestimate the child in relation to the rest of the sample. But two other mothers, similarly identified, did not have notably high FWSs in relation to ELA.

CONCLUSION

Maternal vocabulary scores for ' First Words ' used were found to be significantly and highly correlated to professional assessments of expressive and receptive language ability using the Reynell Developmental Language Scales. No effect for social class was found. These results are subject to two qualifications. Firstly, the mothers in the sample had been involved in an early intervention programme since the early weeks of the child's life and so were familiar with assessments and practised in making observations on their children. Since a primary objective of parental based intervention programmes is to facilitate parental assessments and observations in order to select appropriate teaching goals and activities, the results may not be representative of mothers who have not been in a similar programme. Indeed, given this aim of programmes, comparison between parent-parent and parent-professional assessments on such instruments may offer a useful method of evaluating specific aspects of the intervention programme.

The second qualification to note is that the First Word Checklist used in the main study was limited to nouns. An expanded First Word checklist which included word combinations and contextual use of language would appear more useful (e.g., adapting Gillham's First Sentences Programme, 1982, or Crystal *et al.*'s (1978) LARSP system).

This limitation was highlighted when Reynell scores were predicted from the first word scores using the regression line. These generally resulted in conservative predictions because, it was argued, the Reynell Scales include grammatical usage. Consequently, it was not possible to clarify whether maternal rankings were consistently higher than professional rankings. It should be noted that, firstly, the ability to structure sentences is often differentially impaired or delayed in ESN(S) children and therefore one might expect vocabulary estimates to be higher than Reynell ELAs irrespective of who made them, and, secondly, the common articulation problems found in children with Down's syndrome may result in an underestimation because his/her expressive language is unintelligible to strangers. Thus mothers' apparent overestimation could equally be interpreted as professional underestimation.

Within the context of the study, it was possible to identify a small number of relatively large discrepancies between maternal-professional rankings.

These discrepancies could indicate:

(i) an unreliable professional assessment. For three discrepancies this was found to be the case; the professional assessment had noted difficulties when testing the child.

(ii) an idiosyncratic language ability in the child, i.e., the child may have an extensive single word vocabulary but no grammatical usage, or that the child's use of language was specifically related to contexts other than that presented in the assessment session.

(iii) that the maternal ratings were unreliable. The mothers may not have sufficient knowledge of the child, or may have difficulty in establishing a realistic appraisal of the child.

Thus where discrepancies are found a number of hypotheses can be generated to account for the sources of variance. More information will be required to explain the variance and this is an important part of the assessment and intervention procedure. In the case of (i) and (ii) above, this extended exploration is necessary if an accurate assessment, and consequently an understanding of the child's needs, is to be established. In the case of (iii) the information is of particular importance in directing the professional guidance/counselling of the parent. Hunt and Paraskevopoulos (1980), for example, noted that mothers who rated their children much higher than the diagnostician appeared to have excessively high expectancies of the child. They argue that these mothers are less likely to provide realistic and attainable teaching occasions during their daily interactions with the child. Since a major aim of parent training programmes is to help parents make realistic assessments, and so optimise their daily interactions, this source of variance is well worth exploring.

The present study was concerned with young children in an intervention programme. Generally parents of such children are reasonably optimistic and have positive expectations. However, such expectations can alter as the child grows older and the degree and consequences of the handicap become more apparent. Bryan *et al.* (1982) report that mothers of learning disabled school age children describe the child more negatively in the areas of current attainment and future performance than do mothers of non-disabled children. They also argue that these expectations and mothers' own feelings of less self-efficacy may be particularly crucial factors in the success of special education programmes.

Thus the regular monitoring of parental expectations would appear to be as important as regular assessments of the child. If parents were involved in the assessment process this would take place. The use of instruments such as FWS may provide a formal way of securing this involvement.

In conclusion, and recognising the limitations of the present study, we would suggest that parents of handicapped children can make reliable assessments of their children's attainments. We do not suggest that mothers' reports should replace standardised assessments (e.g., Reynell) but that the two sources of information are complementary. The child will often talk more in a familiar setting and use more mature language when it serves an important function. The parents' high level of contact gives the opportunity to observe this and the practice to interpret idiosyncratic forms and poor articulations. But like any professional, the parents will require some guidelines if they are to use formal assessment instruments and procedures. Given this, sources of variance between professional and parent assessments should be seen not in terms of parental bias, but as useful information to direct further exploration of the child's abilities and/or guidelines for parent counselling. Provided parental input into the assessment process is seen as having equal status to that of the professional, then discussion of sources of variance should reduce potential friction and make a positive contribution to the child, parent and professional.

ACKNOWLEDGMENTS.—We would like to thank Paul Nicklin, Eileen Oakes and Peter Wilkinson who collected the data as part of their M.Sc. (Educational Psychology) dissertations and Dorothy Jeffree and Peter Mittler for their comments on the draft. The study was part of a larger programme sponsored by the Department of Health and Social Security.

REFERENCES

BAYLEY, N. (1969). *Bayley Scales of Infant Development.* New York: The Psychological Corporation.

BROWN, R. I., and SEMPLE, L. (1970). Effects of unfamiliarity on the overt verbalisation and perceptual motor behaviour of nursery school children. *Br. J. educ. Psychol.*, 40, 291-293.

BRYAN, T., PEARL, R., ZIMMERMAN, D., and MATTHEWS, F. (1982). Mother evaluations of their learning-disabled children. *J. special Educ.*, 16, 149-159.

CUNNINGHAM, C. C. (1983). Early support and intervention: the Hester Adrian Research Centre Infant Project. In MITTLER, P. and MCCONACHIE, H. (Eds), *Parents, Professionals and Mentally Handicapped People: Approaches to Partnership*. London: Croom Helm.

CRYSTAL, D., FLETCHER, P., and GARMAN, M. (1978). *The Grammatical Analysis of Language Disability*. London: Edward Arnold.

DEPARTMENT OF EDUCATION AND SCIENCE (1983). *Assessment and Statements of Special Educational Needs*. Joint circular from DES and DHSS, January.

EWERT, J. C., and GREEN, M.W. (1957). Conditions associated with the mothers' estimates of the ability of her retarded child. *Am. J. ment. Defic.*, 62, 521-533.

FAMILY FOCUS (1982). *Parental Involvement in the Assessment of Children with Special Needs.* A discussion paper by Coventry Parent Group. Family Focus, c/o 9 Fairmile Close, Ernsford Grange, Coventry.

GILLHAM, B. (1979). *The First Word Language Programme.* London and Beaconsfield: Allen and Unwin, and Beaconsfield Publishers.

GILLHAM, B. (1982). *Two Words Together: A First Sentence Language Teaching Programme.* London: Allen and Unwin.

GRADEL, K., THOMPSON, M. S., and SHEEHAN, R. (1981). Parental and professional agreement in early childhood assessment. *Topics in Early Childhood Special Education*, 1, 31-39.

HUNT, J. MCV., and PARASKEVOPOULOS, J. (1980). Children's psychological development as a function of the inaccuracy of their mothers' knowledge of their abilities. *J. genet. Psychol.*, 136, 285-298.

JEFFREE, D. M., CHESELDINE, S. E., and SHORVON, C. (1981). Sources of variance between parents' and professionals' estimates of independence skills of ESN(S) children. *Br. J. mental Subnormality*, 32, 66-73.

NEWSON, E. (1976). Parents as a resource in diagnosis and assessment. In WOODFORD, P., and OPPE, T. (Eds). *Early Management of Handicapping Conditions.* IRMMH Review of Research and Practice, 19. London: Associated Scientific Publishers.

REYNELL, J. (1977). *Reynell Developmental Language Scales* (Revised). Windsor: NFER.

SIMEONSSON, R. J., HUNTINGTON, G. S., and PARSE, S. A. (1980). Assessment of children with severe handicaps: multiple problems-multivariate goals. *J. Assoc. Severely Handicapped*, 5, 55-72.

WERNER, E. G., and BAYLEY, N. (1966). The reliability of Bayley's revised scale of mental and motor development during the first year of life. *Child Dev.*, 37, 39-50.

(*Manuscript received 27th May*, 1983)

Br. J. educ. Psychol., 54, 168-176, 1984

USE OF ABBREVIATED COGNITIVE TESTS IN THE ASSESSMENT OF CHILDREN'S SPECIAL EDUCATIONAL NEEDS

By C. J. PHILLIPS
(Centre for Child Study, University of Birmingham)

SUMMARY. The responsibilities of educational psychologists for the assessment of children's special needs have been extended by recent legislation. In seeking ways to meet the extra demands on limited resources, school psychological services may consider more economical ways of obtaining information on the abilities of individuals. For the assessment of cognitive functioning, usually part of diagnosis, shortened forms of intelligence scales, such as Wechsler's, are often favoured. IQs from the full WISC and pro-rated IQs from four subtests are examined for two samples of 10-year-olds. It is shown that the regression on the short-form is significantly different for the 'clinic' population, from the parameters of the general normative population. Abbreviated scales of this kind cannot be recommended for the assessment of special educational needs.

INTRODUCTION

THE law—the *Education Act 1981* and the *Education (Special Educational Needs) Regulations 1983*—requires local education authorities to identify pupils who have special educational needs and to make provisions for them. For some such children the issue of a Statement of Special Educational Needs follows a formal process which must include assessment. The law sets out in greater detail than any earlier enactment how assessment should be conducted. Minimally, the LEA " must seek educational, medical and psychological advice ", which " must be written advice on the features of the case which affect the child's special educational needs and on the ways of meeting these needs " (DES, 1983, para. 21). Psychological advice " must be sought from an educational psychologist employed by the LEA or engaged by the LEA for the purpose of any particular case. . . . Where other psychologists have relevant knowledge or information . . . they must be consulted " (para. 30).

Whatever the disappointments of some people in the new legislation (Newell, 1983), psychologists will feel satisfaction in the recognition that has at long last been given to their distinctive contribution. Less than 20 years ago, children could be sent to special school, or excluded from school altogether, on the evidence of a mental test applied by a school medical officer. But there has been a time lag and many LEA psychologists have now to adjust to the new responsibilities which their predecessors sought. Moreover, the specifications of the law may require some modification of traditional values and practices.

A stated aim is to bring together " the skills, perceptions and insights of professionals in different disciplines. . . . Each professional adviser needs to be aware of the roles of his colleagues and should seek to reach agreement with them on their several roles and functions " (para. 34). Likewise, there is a potential source of conflict between the educational psychologist's obligations to his employing authority and his clinical commitment to the client. Such problems are not new, of course, but there is now a categorical imperative, making extra demands upon existing services.

There is greater openness in the assessment procedures and wider critical scrutiny than before. For example, advice given to the LEA, including the psychological evidence, must be presented to the child's parent when a Statement is issued; the parent may request to see the advice when the issue of a Statement does not follow assessment. Decisions are subject to appeal, at which the parent may offer independent expert evidence.

The quality of professional advice is expressed in general principles: it " should include enough information to explain fully the nature of the case and to lead to a proper understanding of the child's needs " (para. 7). In psychological assessment the " focus should be on the child himself rather than on his disability "; assessment is to be very broadly conceived, to include " the personal resources and attributes of the child "; " his abilities as well as his disabilities " (para. 3). It is stressed that " assessment is a continuous process " (para. 5), not assignment to a status that pertains throughout a child's school life.

Experience of operating under the new legislation may make explicit, what seems to be already implicit, that psychologists' advice must be based upon avowed sources of evidence—first-hand observations of and interviews with the client—that are distinctively psychological and that have known validity, reliability, and relevance. The Circular 1/83 (DES, 1983) contains an Annex which suggests main headings—abilities, attainments, personality factors, and attitudes—of assessment; although this ' Suggested Checklist ' does not form part of the legislation, it will surely be cited as a minimum standard.

The processes of decision-making and the psychological contribution, briefly outlined here, do not apply to the many children with special educational needs which are not " such as to require LEAs to determine special educational provision ". One might expect parents and teachers in the future increasingly to demand that the LEAs interpret their responsibilities for those children to include the same standards of psychological assessment (para. 8-11).

The expectations of clients—children, their parents, and their teachers—which the law encourages, and the rights which it gives them, demand a degree of professional conformity and consistency among educational psychologists to which hitherto they have not needed to aspire, at a time when the profession displays greater diversity than in the past. Problems must arise as to how this professional outlook is to develop and how these responsibilities are to be met, without serious loss to other applications of psychology which have flourished in recent years.

" The reason for doing an assessment should and usually does dictate the form it should take. . . . In general there are three types of assessment required in service contexts . . . (they are) administratively oriented assessment, assessment for programme placement, and individual functional assessment. The first level provides a basis for an administrative decision as to what type of facility will be most suitable for the child " (Kiernan and Jones, 1980). Distinction between this level and the second, programme placement, is not always clear in the assessment of the whole range and variety of special educational needs; it is blurred in the Annex of Circular 1/83. However, the advice to be given by educational psychologists to LEAs must indubitably arise from an assessment which includes the " administratively oriented " level, and which relates " the child's performance to that of other children and is consequently typically norm-referenced . . . The most obvious examples . . . are the IQ and DQ tests ". The need is for instruments with " high validity since decisions . . . may well carry great weight in determining the child's future ". The same authors express the widespread belief in a need for instruments at this assessment level that take as little time as possible to administer. Doubtless much consideration will now be given to economical ways of obtaining information, in framing advice concerning individual special educational needs.

Abbreviated tests of intelligence

Abbreviated forms of psychodiagnostic tests are proposed in order to economise the resources of agencies that employ such instruments. Methodological issues in the use of short forms (SF) of clinical tests were reviewed by Levy (1968), who placed these investigatory tactics within a " decision-theoretic framework ". A basic issue is that of defining the appropriate population for investigation of the characteristics of measurement by SFs.

Primary information on present abilities of children with special educational needs will usually include normative-referenced evaluation of cognitive functioning, one of the items in the ' Suggested Checklist ' (DES, 1983). Certain instruments for the measurement of intellective development are the most eligible for shortening because the technical literature exists. In some of these, shortened forms take smaller samples within a task category, for example, by administering only the starred items of a mental-age level in the Stanford Binet Intelligence Scale (Terman and Merrill, 1961, p. 83). Another basis of selection, that which is the subject of this paper, gives priority to certain kinds of tasks in a battery of subtests. It is illustrated here by reference to the *Wechsler Intelligence Scale for Children* (WISC) because results of testing samples of two populations were available to the author.

Important works dealing with the psychological assessment of children recommend use of abbreviated forms of the WISC, to enable the clinician or researcher to devote more time to other lines of enquiry; for example, Savage (1968, p. 36) and Chazan *et al.* (1974, p. 119). The topic was included in the review of the WISC by Glasser and Zimmerman (1967) and of the WISC-R by Kaufman (1979). Indeed, the output of publications dealing with SFs of the Wechsler Scales is so great that one speculates on the balance of labour expended to labour saved in not giving every WISC testee the full battery (FS).

Development of short forms arose " from a desire to curtail testing time . . . Some researchers, . . . more impressed by the global IQ than by differential abilities, . . . advocate the use of abbreviated versions . . . to make important educational decisions such as special education placement. . . . Should abbreviated WISC-Rs be avoided altogether in clinical assessment? Ideally, yes. Practically, there are times when screening is necessary or desirable " (Kaufman, 1979, p. 206), but that author does not examine what is at stake in following the path of expediency. Cronbach had earlier observed that " in case work . . . (the WISC SF) is not a proper basis for serious decision ", but it " may properly be used in research in which conclusions are reached by comparing groups of subjects " (1970, p. 225). Selected WISC subtests can be instruments in their own right, without recourse to the claim that they are estimates of WISC IQs; if, however, that claim is made, the same difficulty arises with research subjects as with clinic cases.

Various criteria have been proposed for the selection of a few WISC subtests. Kaufman, in recommending four tests, Arithmetic, Vocabulary, Picture Arrangement, and Block Design, had regard to " the degree to which the tasks are good representatives of their own Scales (Verbal or Performance) rather than the Full Scale ". The use of four tests has been most often advocated in this country: Similarities, Vocabulary, Block Design, and Object Assembly. " The sum of all four subtest scale scores provide a reliable estimate of Full Scale IQ " (Rutter *et al.* 1970, p. 28). The origin of that selection was Maxwell's factor analysis (1959) showing those tests to be the best measures of the ' verbal intellectual ' and the ' space performance ' dimensions which he abstracted from Wechsler's intercorrelation matrix, and which he dissociated from the Verbal and Performance Scales.

There is a basic problem inherent in the decision to limit testing to certain tasks within the WISC and similar instruments. It is not a problem of psychometric reliability: in some circumstances, it is sensible to make do with less, and therefore less reliable, evidence. The use of particular subtests is not in itself the problem; it becomes a problem when scores from a few prescribed tasks are claimed to be valid estimates of WISC IQs, or of Verbal or Performance quotients, and to provide information for an authoritative judgment about a child's overall cognitive abilities.

Evaluation of WISC SFs are made from samples of the general child population; most commonly, the selection of tests and determination of their weightings will have been based upon parameters of that population. But instruments like the WISC are not used for classification or selection of children in general; they are used clinically, or, when they are employed in surveys, most often in the study of children who are educationally backward or behaviourally deviant, as in the Isle of Wight study (Rutter *et al.*, 1970). Therefore, the important population for which one needs to know the standing of a WISC SF is children referred to school psychological and guidance services by teachers, parents, and others, who have grounds for concern about learning or behaviour, when the clinician's initial objective is a diagnosis of the nature of the problem and he judges cognitive functioning to be relevant to that diagnostic process. One would extend that definition to include not only the formally referred, but also those detected in other contexts. Putting aside all other factors bearing on his decision, can the clinician be confident that the ' referral ' population is like the normative population with respect to the parameters which determine WISC SFs?

In discussing the use of selected subtests to derive WISC IQs, " in order to lighten the burden of testing ", Cronbach (1970) judged that correlation between SF quotients from two or more subtests and FS quotients " is fallacious evidence because the two measurements are not independent ". The question, " how adequately does the part score correlate with the total score it replaces? ", had been answered by Levy (1967): satisfactory answers depend, among other things, upon the population of study. " The regression slope of the FS scores on SF scores provides a more appropriate test of whether or not a SF developed on a wide range sample remains satisfactory for more restricted groups " such as the clinic population (Levy, 1968). We are to examine some SF results by this criterion.

METHOD

Results from administration of the WISC were obtained from two samples of English children, aged between 10 years 4 months and 10 years 7 months, i.e., the same as Wechsler's 10½-year-olds (1949, Table V). They were given 10 subtests of the Scale, not including Digit Span and Mazes. From the scaled subtest scores, Verbal, Performance, and Full Scale IQs were obtained as the Manual prescribes. The samples will be denoted ' General ' and ' Clinic '.

The General sample consisted of 200 children, with equal numbers of boys and girls, to represent proportionately the socio-economic groups by fathers' occupations (B.P.S., 1955; Phillips, 1983). The estimated mean Full Scale IQ of 10½-year-old English children was 103, with a standard deviation of 16.

The Clinic sample was obtained by an approach to a number of child guidance clinics and school psychological services for results from WISC protocols of children examined in the 1960s. They were to be children of the same age as the General sample, who had been given initial psychological assessments for problems of learning and/or adjustment, not children tested in the course of a survey nor those requiring only a review of educational placement. The testees were to be indigenous, not immigrants with problems of mixed language in the family or from unusual cultural backgrounds. It was desired to exclude cases in which it was recognised at the time of testing that rapport had not been established. Only examinations carried out by an experienced psychologist were to be included. The sample obtained (N = 65) was smaller than planned; it comprised all the cases found in the clinical records of six institutions including the Centre for Child Study, University of Birmingham. A few of those approached would not co-operate because it was felt that it would involve a breach of confidentiality, although identities were not needed, or that it would be too difficult to obtain the permission of medical directors or local authorities. But the main reason for the shortfall of cases was the requirement that all 10 subtests of the WISC had been administered. It is not known how truly representative is this sample of the clinical-referral population. The range of FS IQs was 45 to 119; there were about twice as many boys as girls.

Calculations of SF quotients

A major limitation of this study, as of most others, is that it is retrospective: records of subjects given the full scale have been re-scored for the short forms. That is, one is assuming that an independently administered short form will behave like the selected subtests embedded in the full battery (Levy, 1968).

The two English samples are used to examine the most favoured abbreviation of the WISC: two subtests, Similarities and Vocabulary, estimate Verbal IQ; two subtests, Block Design and Object Assembly, provide a short-form Performance IQ; from the same four subtests combined the FS IQ is estimated.

The most often employed method for deriving quotients from two or three subtests is by pro-rating, in the way Wechsler provides for pro-rating scaled scores from four subtests in Appendix F of the Manual. The pro-rated (P-R) VQ, PQ, and FSQ values reported in Table 1 are obtained by multiplying the sums of scaled scores by 2·5. Another method, rarely used except in research, is via multiple regression equations. The formulae here used, based upon statistics of the General sample, are:

$$\text{VQ} = 53{\cdot}7+(1{\cdot}7^{*}\ \text{Similarities})+(2{\cdot}9^{*}\ \text{Vocabulary})$$

$$\text{PQ} = 62{\cdot}3+(2{\cdot}7^{*}\ \text{Block Design})+(1{\cdot}0^{*}\ \text{Object Assembly})$$

$$\text{FSQ} = 48{\cdot}5+(0{\cdot}9^{*}\ \text{Sim.})+(2{\cdot}7^{*}\ \text{Voc.})+(1{\cdot}2^{*}\ \text{B.D.})+(0{\cdot}3^{*}\ \text{O.A.})$$

Shortform quotients calculated by this method (M-R) are included in Table 1 merely for comparison with pro-rated quotients (P-R).

RESULTS

In Table 1 are given the results of testing that " the slope of the two regressions are the same " (Snedecor, 1956, pp. 397-398). Comparisons of the regression slopes of the two samples are sharply contrasted in the Verbal and Performance scales, with

TABLE 1

WISC SF QUOTIENTS, MEANS AND REGRESSION COEFFICIENTS FOR GENERAL AND CLINIC SAMPLES

	General sample (N = 200)			Clinic sample (N = 65)		Significance of difference between regression coefficients
	Mean (SD)	SE (5%)	*b*	Mean (SD)	*b*	F-ratio (df 1,261)
Verbal scale						
IQ (FS)	104·3 (15·6)			90·7 (15·9)		
SF (P-R)	105·1 (14·6)	11·3	0·99	95·3 (17·7)	0·83	10·78
SF (M-R)	103·6 (10·7)	11·2	1·34	96·3 (12·8)	1·13	10·69
Performance scale						
IQ (FS)	102·8 (15·3)			91·7 (17·4)		
SF (P-R)	105·2 (19·4)	14·9	0·69	94·9 (22·2)	0·69	0·01
SF (M-R)	103·5 (10·6)	16·3	1·22	97·4 (12·1)	1·26	0·18
Full scale						
IQ (FS)	104·0 (15·8)			90·3 (16·8)		
SF (P-R)	105·7 (16·4)	11·7	0·89	94·7 (18·8)	0·83	1·92
SF (M-R)	104·4 (11·4)	11·4	1·29	96·1 (13·3)	1·16	3·53

unambiguous results. Pro-rating the two subtests, Block Design and Object Assembly, to form an estimate of Performance IQ works in the Clinic sample as well as it does in the General sample. But regression of Verbal IQ on the pro-rated sum from Similarities and Vocabulary is very significantly different ($P < 0{\cdot}001$) for the two samples. One can have no confidence, from these results, that the SF method of assessing verbal cognitive abilities may be employed within the clinic population.

Since Full Scale IQ is a blending of Verbal and Performance, one faces the more difficult issue of evaluating SF Full Scale IQ. Significance of difference between regression slopes of the two samples does not reach conventional probabilities; see Table 1. The values from SE of Estimate within the General sample, at 5 per cent level, are shown in Table 1. From the reliability coefficients of the three WISC scales, at age 10½ years, given in the Manual (Table VII), one may calculate the standard error of measurement at the 1 per cent level of probability: 7·7, 12·8, and 8·6 for Verbal, Performance, and Full Scale IQs respectively. These values are adopted to form a confidence zone (Nunnally, 1978, pp. 239-241), in order to compare the two samples for the occurrence of SF quotients falling outside the zone.

As shown in Table 1, a general feature of SF IQs is overestimation; furthermore, since the average quotients of the Clinic sample are below the estimated means of the population, the regression effect will produce more cases of overestimation in that sample than in the General sample. In Table 2, therefore, percentages are given for significant over- and underestimation of estimated ' true ' score when SF quotients are used in the Clinic and General samples. Consistent with the similarity of regression slopes, percentages for the Performance Scale are comparable in the samples. As expected, percentages for the Verbal Scale are very different. The Full Scale values are intermediate; they show 25 per cent of clinic cases to be wrongly estimated by a significant margin ($P < 0{\cdot}01$) when the short form of the test is used. The reader must decide whether it would have been an unacceptable risk to assess the Full Scale IQ of these clinic cases from the four WISC subtests. There surely can be no doubt that it would have been unsound to make inferences about ' verbal-performance ' discrepancies by such abbreviated WISC testing.

TABLE 2

PRO-RATED (P-R) WISC SF QUOTIENTS, PERCENTAGES OF SIGNIFICANT OVER- AND UNDER-ESTIMATION OF ESTIMATED ' TRUE ' IQs

	General sample (N = 200)	Clinic sample (N = 65)
Verbal scale		
Over by > 7·7 points	10·5	26·2
Under by > 7·7 points	5·0	3·1
Performance scale		
Over by > 12·8 points	17·0	15·4
Under by > 12·8 points	7·0	3·1
Full scale		
Over by > 8·6 points	14·0	23·1
Under by > 8·6 points	5·0	1·5

There have been a few studies that attempt solutions to this problem. For example, Mumpower (1964) examined WISC records of " 50 exceptional children for evaluation "; a short form based on the Vocabulary and Block Design subtests would have led to misclassification of 22 per cent of his cases.

DISCUSSION

The WISC is a test of general intelligence. " Its author believes that . . . it is possible to measure it objectively; and that, by so doing, one can obtain a meaningful and useful index of a subject's mental capacity " (Wechsler, 1974). His rationale of assessment is that of sampling a number of different tasks, which Wechsler likens to different languages; these " may be easier or harder for different subjects, but it cannot be assumed that one language is necessarily more valid than another ". These terms do not admit of some WISC subtests being better measures of IQ than others. If one does not accept Wechsler's terms, one may propose a different plan for combining subtest scores, as did Maxwell (1959) in recommending two factor-based measures, but he was not then offering an abbreviated form of the Scale.

For those who accepted Wechsler's premises, estimation of an individual's WISC IQ by administration of fewer than the standard ten subtests was thought to be an option. Which subtests? How to choose them? Answers to the questions were sought from evidence of the subtest scores on the full scale by children in general, at particular ages, from observing that certain tasks have higher factor loadings on a dimension, or that they are the most potent variables in multiple prediction. The statistical procedures, which involve averaging relationships among the subtest variables, can provide a best estimate of IQ from a restricted sample of tasks for randomly selected members of the population. The SF process does not sanction an invariable preference for those tasks in the assessment of an individual, especially one who belongs to the heterogeneous population of deviance and disability, learning difficulty, and emotional disturbance, which requires psychological investigation. In many instances, the impulse to seek psychological aid on their behalf arises from awareness of exceptional disparities in things that normally go together.

Some educational/child psychologists hold doubts about the concept of general intelligence or have little confidence in its measurement through a set of prescribed tasks. But the expectations of other professionals with responsibilities in the field, the demands of institutions, or the exigencies of making special educational provisions, may lead to reporting information of IQ, grudgingly or with reservations. Perhaps the worst possible tactic in that dilemma would be the routine use of abbreviated scales.

The Wechsler Intelligence Scale for Children was superseded by the WISC-R (Wechsler, 1974). These results, obtained some years ago, do not necessarily apply to the revised test, although it would be surprising if substantial differences were found. Other scales of cognition present similar problems in clinical use. In the British Ability Scales four subtests, at different age levels, are prescribed for obtaining a SF IQ. But its authors, exceptionally, do so within a context which stresses the misleading character of the IQ, General or Short-form, when there are marked variations in the levels of an individual's tested abilities, and which espouses the principle of " ipsative measurement: the identification of strengths and deficits in ability within the individual " (Elliott, 1983, p. 136).

The obverse of the SF issue has greater theoretical and practical ramifications. Indices like intelligence quotient, mental age, and language age, are usually composites from a number of tasks. For a child with special educational needs, they may be an average of abilities that are exceedingly disparate. For example, what *can* it mean to report of an 11-year-old with WISC VQ 128 and PQ 69, that he is a boy of average abilities, because his FS is 100? Or, what is communicated in an ITPA Language Age of 4 years, for a young child functioning at the 5-year-level in the visual-motor channel, whose performance in the ITPA auditory-vocal subtests did not exceed what is typical of 3-year-olds?

Some academic acquaintances view these arguments and findings as having no relevance to current practice. Psychologists may indeed " no longer find it so useful to give individual, norm-referenced tests of hypothetical constructs " (Gillham, 1978, p. 84) but the writer's impression is of a field of applied psychology which in its diversity includes traditional-type mental measurement, while tending increasingly to neglect to give attention to its problems and pitfalls. The use of such tests continues to be a major feature of most educational psychologists' practice. Seventy-one per cent of the children referred for investigation in Portsmouth in 1975 were administered an intelligence test (Wright and Payne, 1979). A questionnaire completed in 1978 by about one-third of all practising educational psychologists in England and Wales showed that " the WISC is clearly the most used test "; 82 per cent of the respondents regularly used it and a further 12 per cent occasionally did so (Quicke, 1982). In the field of child care and delinquency, where one might not expect intelligence testing to be a primary concern, Moss and Sutton (1981) found that educational psychologists of a large West Country authority " administered all or part of an intelligence test, usually the WISC " in *all* cases of the type seen over a 12-month period.

The writer agrees wholeheartedly that " the assumptions of classical test theory no longer provide an adequate framework for a systematic approach to planned intervention " (Eggert, 1982). But decisions have to be made about needs, probable syndromes, and the like. It is here that " some practitioners, particularly in the sphere of school administration, have misused psychometric test data (either by administering a single subtest from a test battery, or by combining subtest scores to give a meaningless IQ-equivalent) in order to classify and allocate children ".

If use of standardised tests is to continue as a substantial part of the practice of educational and clinical psychologists, partly as a consequence of the 1981 Act, and if such psychological measurement constitutes part of the operational definitions of ' learning difficulty ' and ' special educational need ', the theoretical issues and the technology demand much more serious attention than they have received in recent times in this country.

ACKNOWLEDGMENTS.—Thanks are due to Beryl Smith, Ronald Gulliford, and Andrew Sutton, erstwhile colleagues, for their constructive comments.

REFERENCES

BRITISH PSYCHOLOGICAL SOCIETY, COMMITTEE OF PROFESSIONAL PSYCHOLOGISTS (MENTAL HEALTH) (ENGLISH DIVISION) (1955). *Final Report of the Working Party on the WISC.*

CHAZAN, M., MOORE, T., WILLIAMS, P., and WRIGHT, J. (1974). *The Practice of Educational Psychology.* London: Longman.

CRONBACH, L. J. (1970). *Essentials of Psychological Testing* (3rd edit.). New York: Harper Row.

DEPARTMENT OF EDUCATION AND SCIENCE (1983). *Circular* 1/83, *Assessments and Statements of Special Educational Needs.* London: HMSO.

EGGERT, D. (1982). Family, environmental and developmental variables in mental retardation: a multi-dimensional approach. *Except. Child,* 29, 87-99.

ELLIOTT, C. D. (1983). *British Ability Scales, Manual* 1, *Introductory Handbook.* Windsor: NFER-Nelson.

GILLHAM, W. E. C. (Ed.), (1978). *Reconstructing Educational Psychology.* London: Croom Helm.

GLASSER, A. J., and ZIMMERMAN, I. L. (1967). *Clinical Interpretation of the WISC.* New York: Grune Stratton.

KAUFMAN, A. S. (1979). *Intelligent Testing with the WISC-R.* New York: Wiley.

KIERNAN, C. C., and JONES, M. (1980). The Behaviour Assessment Battery for use with the profoundly retarded. In HOGG, J. and MITTLER, P. J. (Eds), *Advances in Mental Handicap Research,* Volume 1. Chichester: Wiley.

LEVY, P. (1967). The correction for spurious correlation in the evaluation of short-form tests. *J. clin. Psychol.,* 23, 84-86.

LEVY, P. (1968). Short-form tests: a methodological review. *Psychol. Bull.,* 69, 410-16.

MAXWELL, A. E. (1959). A factor analysis of the Wechsler Intelligence Scale for Children. *Br. J. educ. Psychol.,* 29, 237-241.

MOSS, G., and SUTTON, A. (1981). Educational psychologists and the juvenile court. In LLOYD-BOSTOCK, S. (Ed.), *Law and Psychology*. Oxford: Centre for Socio-Legal Studies.

MUMPOWER, D. L. (1964). The fallacy of the short form. *J. clin. Psychol.*, 20, 111-113.

NEWELL, P. (1983). *ACE Special Education Handbook—the new law on children with special needs.* London: Advisory Centre for Education.

NUNNALLY, J. C. (1978). *Psychometric Theory* (2nd edit.). New York: McGraw-Hill.

PHILLIPS, C. J. (1983). Wechsler Intelligence Scale for Children: A historical note on its use with English children at 10½ years. *Occasional Papers of the D.E.C.P., British Psychological Society*, 7, No. 1.

QUICKE, J. C. (1982). *The Cautious Expert: A Social Analysis of the Developments in the Practice of Educational Psychology.* Milton Keynes: Open University Press.

RUTTER, M., TIZARD, J., and WHITMORE, K. (Eds) (1970). *Education, Health and Behaviour.* London: Longman.

TERMAN, L. M., and MERRILL, M. A. (1961). *Stanford-Binet Intelligence Scale, Manual for the Third Revision Form L-M.* London: Harrap.

SAVAGE, R. D. (1968). *Psychometric Assessment of the Individual Child.* Harmondsworth: Penguin.

SNEDECOR, G. W. (1956). *Statistical Methods Applied to Experiments in Agriculture and Biology* (5th edit.). Iowa State University Press.

WECHSLER, D. (1949). *Wechsler Intelligence Scale for Children.* New York: Psychological Corporation.

WECHSLER, D. (1974). *Wechsler Intelligence Scale for Children—Revised.* Windsor: NFER.

WRIGHT, J., and PAYNE, T. (1979). *An Evaluation of a School Psychological Service: The Portsmouth Pattern.* Winchester: Hampshire County Council.

(*Manuscript received 29th June*, 1983)

Br. J. educ. Psychol., 54, 177-188, 1984

CONSISTENCY OF CAUSAL ATTRIBUTIONS FOR EXPECTED AND ACTUAL EXAMINATION OUTCOME: A STUDY OF THE EXPECTANCY CONFIRMATION AND EGOTISM MODELS

By J. W. CHAPMAN AND MERYL M. LAWES
(*Massey University, New Zealand*)

SUMMARY. The expectancy confirmation and egotism models of causal attributions for success and failure were investigated with 400 fifth form students taking the New Zealand School Certificate English examination. Attributions for expected exam outcome and actual exam outcome were assessed with a scale containing 33 items categorised into the attribution dimensions of stability and locus of causality. The results suggested that exam outcome attributions are a function of both expectancy confirmation and valence of actual outcome, with the effects being due primarily to subjects whose expectations for passing were disconfirmed by actual failure. Compared to pre-exam attributions based on expected outcome, these disconfirmed failure subjects reduced post-exam attributions to internal (stable and unstable) and external stable causes, but increased their attributions to external unstable causes. Where expectations were confirmed (pass or fail) attributions from pre- to post-exam were generally consistent; however, the confirmed pass subjects tended to see internal factors as more important than the other subjects, and external unstable factors as least important. The findings are discussed in terms of the implications for future achievement situations.

INTRODUCTION

CURRENT interest in attribution theory has generated numerous studies on the relationship between causal perceptions of success and failure and achievement-related outcomes (Bar-Tal, 1978; Weiner, 1979, 1980; Kelly and Michela, 1980). At a general level, it is suggested that individuals usually attribute success to internal factors, such as ability and effort, whereas failure is ascribed to external factors such as task difficulty, lack of assistance from teachers, parents or peers, and bad luck (e.g., Frieze and Weiner, 1971; Nicholls, 1975; Miller, 1976; Arkin and Maruyama, 1979; Davis and Stephan, 1980).

Various exceptions to this pattern of success/failure attributions, however, have become evident in the research literature. Bar-Tal and Darom (1979), for example, found that elementary school pupils tended to attribute success to external causes, and failure to internal causes. Similarly, studies of low achieving or failure-prone individuals have reported the tendency for success to be seen as externally caused and failure as internally caused (Chapman and Boersma, 1979; Pearl *et al.*, 1980). In addition, sex differences in success/failure attributions have also been noted. Some studies (Feather, 1969; Simon and Feather, 1973; Bar-Tal and Frieze, 1977; Deaux and Harris, 1977; Frieze and Bar-Tal, 1980) suggest that females more than males make external attributions for success *and* failure. Other studies (Dweck and Repucci, 1973; Nicholls, 1975; Dweck *et al.*, 1978) found that females made more attributions to external factors for success and internal factors for failure than males.

Such apparent inconsistencies in attribution patterns for school-related task outcomes have been explained in terms of expectancy confirmation (Frieze and Weiner, 1971; McMahan, 1973; Valle and Frieze, 1976; Weiner, 1979). In this model, causal attributions following success or failure are seen as a function of individual expectations regarding performance outcomes. Thus, expected outcomes, either success or failure, tend to be attributed to stable factors, whereas unexpected outcomes tend to be attributed to unstable factors (Feather, 1969; Frieze and Weiner, 1971; Feather and Simon, 1973; McMahan, 1973; Valle and Frieze, 1976). Concomitantly, the greater the discrepancy between expected and actual outcome, the stronger the attribution to unstable factors and the smaller the discrepancy, the greater the attribution to stable factors.

Davis and Stephan (1980) take issue with this expectancy confirmation approach. They assert that the valence of the actual outcome (success or failure) is the primary determinant of post-performance attributions. Thus, attributions to internal factors (ability, effort) will increase with the magnitude of success in the outcome, whereas attributions to external factors (task difficulty, luck) will increase with the degree of failure. Davis and Stephan refer to this interpretation in terms of egotism, whereby failure is ascribed to external factors for which the individual cannot be held responsible (ego-defensive), and success is ascribed to internal factors (ego-enhancement).

In support of this position, Davis and Stephan (1980) present data from a study with undergraduate university students. They found that the students generally had high expectations for success on a class exam, and that these were associated with ability and effort attributions. Actual exam outcomes consistent with those high expectations were again attributed internally to ability and effort, whereas low performance was attributed externally to test difficulty. However, in a regression analysis of the place of ability, effort, test difficulty and luck in predicting the difference score between expected result and actual result, Davis and Stephan found little or no relationship between those causal factors and degree of expectancy confirmation. Thus, it was concluded that the egotism model emphasising locus of causality was supported, whereas the confirmation model based on stability of causality was not supported.

It is debatable whether attributing failure to test difficulty represents an external *and* stable cause, as implied by Davis and Stephan. Meyer (1980) suggests that when test difficulty refers to a university exam it should be considered an unstable factor. This viewpoint seems tenable given that university exams are seldom standardised, fixed entities of known property in the manner of many commercially distributed college entrance or selection exams (e.g., College Entrance Examination Board series, Scholastic Aptitude Test, Graduate Record Examinations). Reconsidering test difficulty in the Davis and Stephan study as an external unstable cause then, provides support for the expectancy confirmation model.

Similarly, it is questionable whether attributions to effort are unstable, as implied by Davis and Stephan and numerous other writers (e.g., Frieze and Weiner, 1971; McMahan, 1973; Bar-Tal, 1978). Indeed, Cooper and Burger (1980), in their synthesis of coding schemes for treating free-response attributions, differentiate between " typical effort " (stable) and " immediate effort " (unstable) attributions. Reinterpretation of Davis and Stephan's findings that expected success was associated with attributions to ability *and* effort, where effort is supposedly exerted typically as a matter of common practice in situations of high valence and ego involvement, indicates success ascriptions to internal stable causes. Unexpected failure on the other hand may be attributed to the external *unstable* factor of test difficulty. Clearly, Davis and Stephan's data do not necessarily support the egotism model over the expectancy confirmation model.

In view of the confusion in the literature over the explanatory value of the egotism and expectancy confirmation approaches, it seems important to clarify if they are indeed alternative viewpoints, or whether they are interactive. A more fruitful approach in considering these competing claims would be to analyse the expectancy and egotism models both separately and in interaction with each other. Accordingly, the main purpose of the present study is to test the egotism valence-outcome model with the expectancy confirmation model in terms of stable/internal, stable/external, unstable/internal and unstable/external causal attributions for success and failure on a major and highly significant school exam.

A second purpose of this study was to investigate the consistency of causal attributions in terms of expected and actual outcomes. Although Davis and Stephan (1980) considered pre-exam attributions based on expected outcome, and post-exam attributions based on actual outcome, their analyses did not allow consideration of the consistency or stability of the pre- and post-causal attributions following expectancy confirmation or disconfirmation. Some studies have reported on attribution consistency for success and failure outcomes across different subject areas (e.g., Bar-Tal *et al.*, 1982) but few, if any, have reported comparisons between attributions for expected and actual outcomes. By analysing pre-exam and post-exam attributions as a function of expected and actual exam outcome, the competing claims of the egotism and expectancy confirmation models may be further tested.

Finally, although there has been a move away from studies of contrived achievement-related events towards the consideration of attributions in " real-life " situations (e.g., Bar-Tal and Darom, 1979; Davis and Stephan, 1980; McMillan and Spratt, 1980; Bar-Tal *et al.*, 1982), there is still a need to consider success and failure attributions for achievement outcomes of high valence. This ensures that the achievement task does engage a strong degree of ego involvement, thereby enhancing the validity of testing the egotism model, as well as maximising the chances that attribution shifts may occur following disconfirmation of expected outcome.

Accordingly, the present study focuses on causal attributions for success and failure in relation to the New Zealand Department of Education national School Certificate English exam. The School Certificate exams are seen as the major form of external subject-based assessment in the New Zealand post-primary school system, and are usually attempted at the end of the fifth form year (at about age 15 to 16 years). Satisfactory passes in a number of subjects are considered as the basic prerequisite for entry into the sixth form, and also as a minimum qualification for many occupations in New Zealand. Further, because English is a compulsory subject for most secondary school pupils, this exam attracts greater interest than most others, with 86·7 per cent of all candidates taking the exam in 1981 (Department of Education, 1982). Clearly the study of causal attributions in terms of expected and actual outcome in School Certificate English has high ecological validity.

In sum, the aim of this study is to investigate the expectancy confirmation and egotism models of causal attributions for success and failure, along with a comparison of expected outcome and actual outcome attributions. The study is located within the context of performance outcome on a highly salient national school exam.

METHOD

Sample selection

A pre-School Certificate English examination attribution questionnaire was completed by 469 Form 5 pupils in November, 1981, within the week prior to sitting the exam. These pupils were enrolled in three North Island co-educational high schools, two located in a provincial city and one in a large rural town. Of this sample, 308 pupils (66 per cent) completed a post-School Certificate English examination attribution questionnaire during the third week of the new academic year in February, 1982. (SC results had been mailed to examinees late in January.) A further 141 post-exam questionnaires were mailed to pupils not returning to school but who had supplied home address information. Completed questionnaires were received from 92 (65 per cent) of these school leavers, with 10 questionnaires returned " address unknown " and 39 non-responses. The final sample therefore comprised 400 subjects, which is 85 per cent of the original sample.

Attribution questionnaire

Pre-examination and post-examination attributions were assessed using a questionnaire developed by the second author (Lawes, 1983). The questionnaire comprised 33 common items of a declarative sentence nature, each of which provided a possible reason for passing/expecting to pass or failing/expecting to fail School Certificate (SC) English. The tense of some items was altered as appropriate according to the context of expected or actual outcomes. On the pre-exam questionnaire, subjects stated whether they expected to pass or fail SC English. They were then required to complete one of two sections, depending on their expectation. The 33 items in each section were the same in terms of content and order of presentation, but worded slightly differently in line with a pass or fail expectation. Subjects rated on a five-point Likert-type scale the extent to which each item was a reason for their stated expected outcome of passing or failing. Similarly, on the post-exam questionnaire, pupils declared whether or not they had passed or failed (a mark of 50 per cent or higher is officially considered a pass) the exam, and rated the same 33 items according to their perceived causal influence on the outcome.

The 33 items in the questionnaire were categorised into the two underlying attribution dimensions originally proposed by Weiner (1972), namely locus of causality (internal/external) and stability (stable/unstable). Using Cooper and Burger's (1980) categorisation scheme for success/failure attributions, nine items were designated stable/internal, ten as stable/external, eight as unstable/internal, and six as unstable/external. The stable/internal causes included items relating to ability, past experience, and ongoing study habits or techniques. For the stable/external dimension, items referred to ongoing help from significant others (mainly teachers and parents), beliefs that the right or wrong questions " always " come up in exams, that there is always an element of luck associated with exam results, and also home study conditions. Unstable/internal causes included exam-related effort, mood on exam, and selection of topics to study for the exam. Finally, unstable/external items dealt with specific exam ease or difficulty, luck on the exam, and marker error.

Examples of reasons for expecting to pass/passing included " Because I have high ability in English "; " Because the English exam will be/was easy "; " Because my parents helped me during the year ". Examples of reasons for expecting to fail/failing included " Because I have low ability in English "; " Because the English exam will be/was hard "; " Because my parents did not help me during the year ".

Internal reliability coefficients (Cronbach's alpha: α) computed for each dimension on the pre- and post-exam questionnaires respectively resulted in the following estimates: stable/internal $\alpha = 0{\cdot}66$, $0{\cdot}77$; stable/external $\alpha = 0{\cdot}71$, $0{\cdot}72$; unstable/internal $\alpha = 0{\cdot}76$, $0{\cdot}82$; unstable/external $\alpha = 0{\cdot}75$, $0{\cdot}73$. The pre and post alpha coefficients for the Full Scale were 0·86 and 0·88 respectively.

RESULTS

Pre- and post-exam attribution scores for each of the locus by stability subscales were calculated by adding the raw scale ratings and dividing by the number of items in the respective subscales. In order to test the competing or interacting effects of the expectancy and egotism attribution models, along with the consistency of attributions before and after the exam, the subscale scores were treated by means of four three-way analyses of variance with repeated measures. The independent variables were expectancy confirmation (confirmed, disconfirmed) and actual outcome (pass, fail), with repeated measures on the pre- and post-exam attributions.

In terms of cell sizes, there were 178 subjects in the Confirmed/Pass group (CP: expected pass, actual pass), 83 in the Confirmed/Fail group (CF: expected fail, actual fail), 23 in the Disconfirmed/Pass group (DP: expected fail, actual pass),

TABLE 1

SUMMARY ANOVA DATA FOR ATTRIBUTION FACTORS

	Internal/ stable		Internal/ unstable		External/ stable		External/ unstable	
	F[a]	p	F	p	F	p	F	p
Confirmation	25·85	0·000	2·20	0·139	4·12	0·043	0·28	0·597
Outcome	17·06	0·000	18·74	0·000	1·49	0·223	8·53	0·004
Pre-post	49·55	0·000	23·83	0·000	6·12	0·014	0·40	0·525
Confirmation × outcome	10·99	0·001	31·54	0·000	2·25	0·114	13·08	0·000
Confirmation × pre-post	10·37	0·001	6·58	0·011	0·56	0·456	3·76	0·053
Outcome × pre-post	23·81	0·000	53·14	0·000	36·03	0·000	2·72	0·100
Confirmation × outcome × pre-post	10·87	0·001	55·18	0·000	8·05	0·005	10·14	0·002

[a] All df's = 1,396.

and 116 in the Disconfirmed/Fail group (DF: expected pass, actual fail). The 201 subjects in the present study who passed the exam represent 50 per cent of the sample, which is in line with the national 50·81 per cent pass rate for examinees in the 1981 SC English exam (Department of Education, 1982).

Analysis of results for the *internal/stable* factor revealed that all main and interaction effects were significant at the 0·001 level or beyond (see Table 1 for summary ANOVA data). An inspection of means showed that higher overall ratings were made on this factor by the Confirmed group than the Disconfirmed group (C $\bar{x}$ = 2·83; D $\bar{x}$ = 2·49), the Pass group compared with the Fail group (P $\bar{x}$ = 2·80; F $\bar{x}$ = 2·52), and on the pre-exam questionnaire than the post-exam questionnaire (Pre $\bar{x}$ = 2·94; Post $\bar{x}$ = 2·59).

Of greater interest, however, are the interaction effects. For the Confirmation × Outcome interaction, the CP group made higher overall ratings on internal/stable items than each of the other groups (CP $\bar{x}$ = 3·08; CF $\bar{x}$ = 2·58; DP $\bar{x}$ = 2·52; DF $\bar{x}$ = 2·46). In terms of the Confirmation × Pre-Post interaction, the Disconfirmed group made a greater change in Pre- to Post-exam ratings than the Confirmed group (Pre-C $\bar{x}$ = 2·92; Post-C $\bar{x}$ = 2·74; Pre-D $\bar{x}$ = 2·84; Post-D $\bar{x}$ = 2·25). Similarly,

TABLE 2

MEANS AND STANDARD DEVIATIONS FOR ATTRIBUTION FACTORS AS A FUNCTION OF CONFIRMATION EXPECTANCY AND ACTUAL OUTCOME

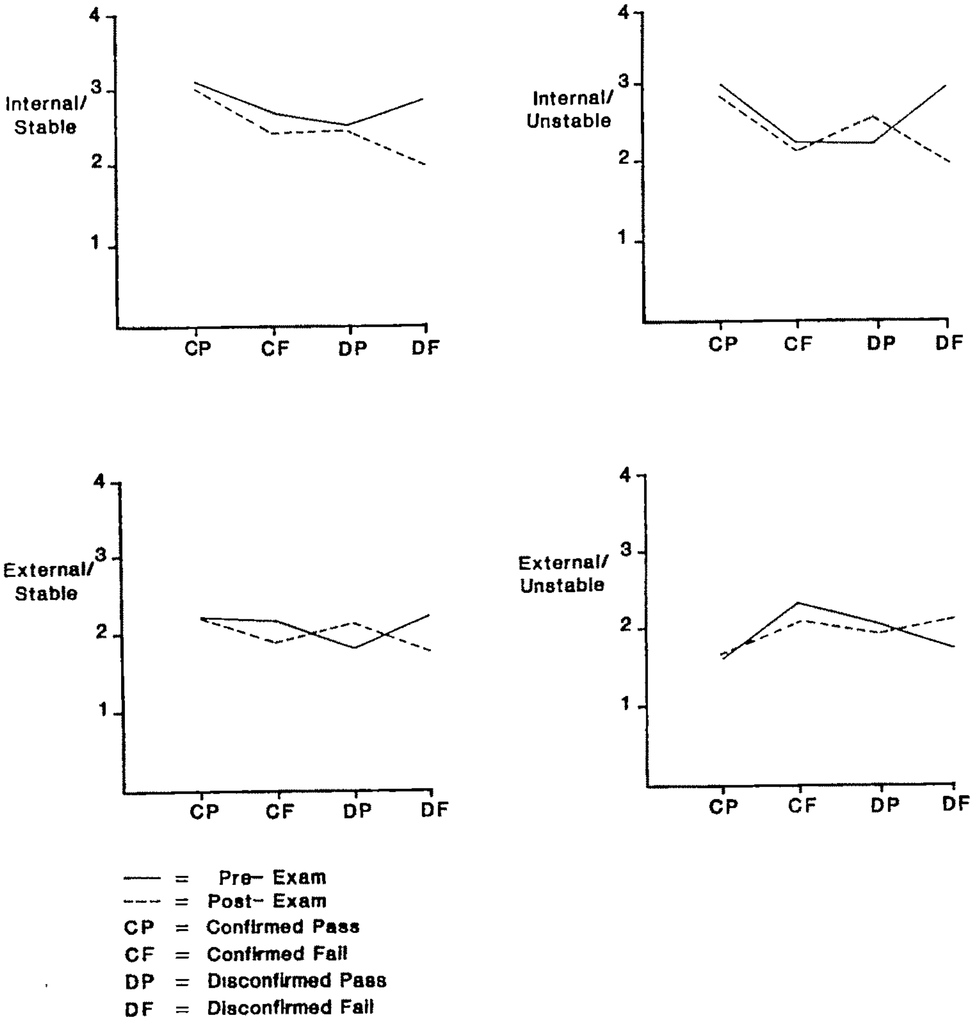
	Confirm pass (N = 178)		Confirm fail (N = 83)		Disconfirm pass (N = 23)		Disconfirm fail (N = 116)	
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
Internal/stable								
Pre-exam	3·13	0·57	2·70	0·65	2·56	0·74	2·89	0·63
Post-exam	3·02	0·57	2·46	0·68	2·47	0·58	2·03	0·63
Internal/unstable								
Pre-exam	3·00	0·64	2·26	0·64	2·25	0·73	3·03	0·64
Post-exam	2·88	0·67	2·16	0·66	2·59	0·67	2·00	0·68
External stable								
Pre-exam	2·21	0·54	2·16	0·71	1·83	0·49	2·21	0·57
Post-exam	2·20	0·53	1·90	0·58	2·12	0·62	2·78	0·55
External/unstable								
Pre-exam	1·68	0·59	2·26	0·75	2·06	0·66	1·77	0·70
Post-exam	1·69	0·61	2·12	0·81	1·95	0·84	2·13	0·78

from the Outcome × Pre-Post interaction it is evident that those who failed the exam made lower ratings following the exam than before (Pre-pass $\bar{x} = 2{\cdot}85$; Post-pass $\bar{x} = 2{\cdot}75$; Pre-fail $\bar{x} = 2{\cdot}80$; Post-fail $\bar{x} = 2{\cdot}25$).

Perhaps the most interesting internal/stable result is the three-way Confirmation × Outcome × Pre-Post interaction. Here the DF group made significant changes in internal/stable ratings following the exam, with the resultant group mean significantly lower than each of the other three groups (minimum F scheffé = 14·41, $P < 0{\cdot}01$). As Figure 1 and the data in Table 2 reveal, the DF group dropped their scores from being 0·24 points lower than the CP group on the pre-test to 0·99 points lower after the exam. Further, that change was significantly greater than the change observed for the CF group (F scheffé = 75·51, $P < 0{\cdot}01$) which also discernibly lowered its internal/stable attributions following the exam.

FIGURE 1

PRE- AND POST-ATTRIBUTION DIMENSION SCORES AS A FUNCTION OF EXPECTANCY CONFIRMATION



Overall, it seems clear that the change in pre-post scores in this group was largely responsible for the other significant interaction effects. Whereas the other groups were relatively consistent in their pre-post attributions, those who expected to pass but failed made a significant post-exam decrease in their internal/stable attributions.

Results for the *internal/unstable* factor reveal that two main effects (Outcome and Pre-Post) and all interaction effects were statistically significant at the 0·01 level or greater (see Table 1). For Outcome, those who passed the exam made higher overall ratings on internal/unstable items than those who failed (P $\bar{x}$ = 2·68; F $\bar{x}$ = 2·36), whereas for the Pre-Post effect, internal/unstable ratings were higher before the exam than after (Pre $\bar{x}$ = 2·81; Post $\bar{x}$ = 2·46).

The Confirmation × Outcome effect was due primarily to the CP group making higher overall ratings on this factor than each of the other groups (CP $\bar{x}$ = 2·94; CF $\bar{x}$ = 2·37; DP $\bar{x}$ = 2·42; DF $\bar{x}$ = 2·51). In terms of the Confirmation × Pre-Post effect, the Disconfirmed group made weaker post-exam ratings of internal/unstable items than each of the other groups (C-pre $\bar{x}$ = 2·63; C-post $\bar{x}$ = 2·52; D-pre $\bar{x}$ = 2·64; D-post $\bar{x}$ = 2·29). Lower post-exam ratings were also shown in the Outcome × Pre-Post interaction for those who failed the exam (P-pre $\bar{x}$ = 2·63; P-post $\bar{x}$ = 2·74; F-pre $\bar{x}$ = 2·64; F-post $\bar{x}$ = 2·08).

Again, the most interesting internal/unstable effect was the three-way Confirmation × Outcome × Pre-Post interaction. Figure 1 and the data in Table 2 indicate that the DF group made similar ratings to the CP group before the exam, probably because both groups expected to pass. These were together significantly higher than the combined CF and DP groups (F scheffé = 24·71, $P<0·01$). After the exam, however, the DF group decreased ratings to internal/unstable items by 1·02 points, putting their attributions in line with the CF group but significantly lower than either the DP or CP post-attributions (minimum F scheffé = 27·14, $P<0·01$). Also of interest is the increase in post-exam ratings by the DP group, to the extent that their internal/unstable scores were significantly higher than either the CF or DF groups (minimum F scheffé = 13·53, $P<0·01$), but nonetheless lower than the CP group (F scheffé = 13·91, $P<0·01$).

To summarise, the effects on the internal/unstable factor were due primarily to a decrease in post-exam attributions by the DF group and an increase in internal/unstable post-attributions by the DP group. Also, both groups that expected to pass the exam (CP, DF) made pre-exam internal/unstable attributions similar to each other; these were higher than the two groups which expected to fail (CF, DP) both of which however also made similar pre-exam ratings. Once again, the combination of expectancy confirmation interacting with actual outcome appears to have the predominant effect on attributions, with disconfirmation, particularly where unexpected failure occurred, resulting in relatively inconsistent pre-post causal ascriptions.

Turning to the *external/stable* factor, Table 1 indicates that the Confirmation and Pre-Post variables produced significant main effects, and the Outcome × Pre-Post and Confirmation × Outcome × Pre-Post interactions were significant.

The main effects were due to the Confirmed group making higher overall ratings than the Disconfirmed group (C $\bar{x}$ = 2·16; D $\bar{x}$ = 1·98), and the pre-exam ratings being higher than the post-exam ratings (Pre $\bar{x}$ = 2·18; Post $\bar{x}$ = 2·01). Those who failed the exam gave rise to the Outcome × Pre-Post effect by making lower external/stable ratings following the exam than before (P-pre $\bar{x}$ = 2·02; F-pre $\bar{x}$ = 2·19; P-post $\bar{x}$ = 2·16; F-post $\bar{x}$ = 1·84). For the Confirmation × Outcome × Pre-Post interaction, the DF group again showed a decrease in post-exam attributions compared to the pre-exam ratings. From making similar ratings to the CP group before the exam, the DF group made significantly lower ratings than the CP group after the exam (F scheffé = 64·41, $P<0·01$). Indeed, the DF post-exam ratings were more in line with the post-exam attributions of the CF group.

In sum, the interaction of expectancy confirmation and actual outcome, particularly in terms of post-exam changes by the DF group, appears to have a predominant influence in explaining results on the external/stable factor.

Finally, the *external/unstable* dimension produced a significant main effect for Outcome, and significant interaction effects for Confirmation × Outcome, and Confirmation × Outcome × Pre-Post. The Outcome effect was due to the Fail group making higher overall ratings on external/unstable items than the Pass group (F $\bar{x}$ = 2·07; P $\bar{x}$ = 1·84). For the Confirmation × Outcome interaction, the significant effect was due mainly to the lower overall external/unstable ratings made by the CP group.

Once again the most interesting result on this factor was in the three-way Confirmation × Outcome × Pre-Post factor. Here the DF and CF groups made the greatest changes in ratings, with the DF group making an increase in external/unstable attributions that was greater in magnitude than the decrease observed for the CF group (F scheffé = 45·57; $P<0·01$). Also of interest is the fact that the DF group moved from being within 0·09 points of the CP group on the pre-exam, and lower than the two groups which expected failure, to being a significant 0·44 points higher than the CP group on the post-exam (F scheffé = 51·22, $P<0·01$).

Overall then, for the unstable/external factor, subjects who failed generally made stronger attributions to these causes than those who passed, whereas those whose pass expectations were confirmed showed the weakest unstable/external attributions. The greatest change occurred for the DF group, who increased their external/unstable attributions following their unexpected failure.

DISCUSSION

The main purpose of this study was to test the competing claims of the expectancy confirmation and egotism models for explaining achievement outcome attributions. A secondary and closely related purpose was to consider the degree of consistency in making attributions prior to an event and after an event, based on expected outcomes and actual outcomes. Both questions are primarily answered by the interaction effects observed in the ANOVA results. In that the presence of interaction effects influences main effects (Hays, 1981), it makes sense in any case to focus on the interactions observed in these findings.

Evidence in support of the expectancy confirmation model may be found in the confirmation by pre-post interaction effects. Here the subjects whose expectations were disconfirmed significantly decreased their internal stable and internal unstable attributions following the exam compared to their pre-exam ratings. The key factor here however appears to be locus rather than stability. The unexpected exam outcome, irrespective of pass or failure, resulted in a lowering of personal responsibility for that outcome. No concomitant increase was found in external attributions. Thus, expectancy disconfirmation leads to the expression of changed attributions following the unexpected result, not along the stable-unstable dimension as found by others (e.g., Frieze and Weiner, 1971; McMahan, 1973; Valle and Frieze, 1976; Meyer, 1980), but instead along the locus dimension.

Support may also be found for the egotism model, which states that the actual outcome is the primary determinant of post-performance attributions. The outcome by pre-post interaction effects revealed that those who failed the exam, irrespective of their initial expectation, made lower attributions to internal stable, internal unstable, and external stable factors after the exam than before. This finding represents a slightly more complex picture where there is some diminished personal responsibility for outcome on both stable and unstable factors, as well as a reduction in attributions to external stable variables such as home study conditions and ongoing teacher and parental assistance. So in decreasing their own responsibility for failure, there is also a reduction in responsibility to fairly constant causes such as conditions for studying and help or lack of it from others during the course of the school year.

While locus is an element in this finding, the result is not as clear-cut as the egotism model holds. That model suggests that success is associated with internal factors and failure with external factors (Miller, 1976; Rosenfield and Stephan, 1978; Davis and Stephan, 1980). The results of this study show that locus is confounded with stability, and that external stable causes are implicated in the reduction of attributions amongst those who failed the exam.

Thus far the picture is not really clear, and neither of the two competing models can claim causal predominance over the other. By turning to the three-way interactions, the results may be more adequately explained. It appears that the findings in this study are due primarily to significant changes in pre-post attributions by the disconfirmed failure group. They made significant decreases in their post-exam attributions to internal stable, internal unstable and external stable causes compared to their pre-exam attributions based on expected successful exam results. Further, they increased their attributions to external unstable causes following the exam over their pre-exam ratings of this factor. These changes in pre-post attributions are more noteworthy when it is considered that the pre-exam attributions for this group tended to be very similar to those of the confirmed pass subjects: both groups expected to pass and made similar attributions according to that expectation. But the unexpected failure led to significant changes in attributions, changes which were more significant than those made by the other subjects whose expectations were disconfirmed in a positive direction. Indeed, the disconfirmed pass group made relatively small changes in their pre-post attributions, with the most noticeable being for a post-exam increase in ratings on internal unstable attributions. This would suggest that factors such as immediate exam-related effort were seen as a variable in causing the unexpected success.

Considering the results overall then, exam outcome attributions appear to be a function of both expectancy confirmation and valence of actual outcome, with the effects being due primarily to subjects whose expectations for passing were disconfirmed by actual failure. Under these circumstances there was a de-emphasis on internal (stable and unstable) and external stable factors following the exam but an increase in attributions to external unstable causes. A corollary to this result is the finding that attributions were generally consistent where expectations were confirmed, despite the outcome, but the direction of attributions was different: the confirmed pass subjects tended to see internal factors as more important than the confirmed fail subjects, and external unstable causes as least important.

Such findings seem logical. Where expectations are confirmed it seems reasonable that attributions are consistent from the pre-exam to the post-exam context. Attributions are made according to an expected outcome and consistently confirmed when the outcome is affirmed. But where expectations are disconfirmed, the direction of the outcome affects the degree of consistency of pre to post attributions, with unexpected failure having the most significant impact. Change in attributions then seem most likely to occur where failure is not anticipated.

Two elements of particular interest emerge from the findings. While it seems consistent with previous studies that those whose failure is unexpected should show a decrease in personal responsibility for the unexpected failure over the anticipated success (Simon and Feather, 1973; Miller and Ross, 1975), it was somewhat surprising to find also a post-exam decrease in attributions by this group to external stable factors. Thus, in addition to reducing their own personal responsibility, there is also relatively little attempt to blame parents or teachers for lack of encouragement and assistance throughout the year, or to attribute the unexpected failure to other external stable causes such as home study conditions of beliefs that outcomes on School Certificate exams in general are largely a function of luck. Instead, by increasing post-exam attributions to external unstable causes such as the difficulty of that particular exam for that year, and of bad luck on that exam specifically, the pupils seem to be implying that ability, effort, and external support exist for utilisation in future exams. In other words, neither they nor their immediate significant others are responsible for the unpredicted failure, and that therefore the possibility of facilitating future successes is not ruled out (Valle and Frieze, 1976; Ostrove, 1978).

Also of interest is the finding that the disconfirmed pass group were not characterised by marked shifts in attributions following the exam. There was certainly not a generalised tendency towards making unstable attributions, as the expectancy confirmation model would suggest. Nor was there a generalised tendency toward making internal attributions following the successful outcome, as the egotism model would suggest. Rather, there were increases in internal unstable and external stable attributions. Thus, exam-specific effort and preparation, along with specific mood states (feeling good during the exam) were seen as relatively important factors. So, too, were the external stable variables of assistance from others and general study conditions. The trend for this group generally then was to doubt their own ability and the usefulness of their study techniques in English, and to express a lack of enjoyment with the subject prior to the exam (internal stable factors) but to see immediate effort and a supportive environment as taking on some importance in causing the unexpected successful outcome. For this group, a sense of confidence in their own ability may have arisen out of the belief that exam revision played a part in the outcome, along with ability and the support of others. As with those whose expectations of success were disconfirmed, the pupils who unexpectedly passed made attributions which imply that successes may be repeated in the future, assuming that exam-related effort is exerted, and significant others continue to be supportive (Valle and Frieze, 1976; Ostrove, 1978).

In conclusion, it appears from this study that causal attributions for success and failure are not as straightforward and predictable as some studies would suggest. In terms of the competing claims made by the egotism and expectancy confirmation models, neither would appear to have causal predominance over the other in exam situations involving high outcome salience. Rather, causal attributions for success and failure appear to be a function of both expectancy confirmation and actual outcome. This effect was relatively specific to those pupils who expected to pass School Certificate English but failed: they made significant changes in their attributions on the basis of their unfulfilled expectations that appeared to affect most if not all of the findings in the present study. The other group of pupils whose expectations were disconfirmed, but who passed, made relatively few changes in their pre-post attributions, while those pupils whose expectations were confirmed were quite consistent in their attributions of expected and actual outcome.

The causal dimensions used to explain expected and actual examination outcome were also more complex than implied in the literature to date. Using just two of the dimensions, stability and locus, it appears that the results of previous studies are not consistent with those obtained under the present conditions of high salience. Thus, unexpected failure produced an increase in attributions to external unstable factors, and a decrease to the other factors of internal stable, internal unstable, and external stable causes. Implicit here is the belief that this unexpected negative outcome in a " one-off " exam was an unusual event with no implications for the development of negative self-esteem or for a lessening of the supportive role played by others. On the other hand, those who unexpectedly passed both took credit for immediate exam-related effort and also gave credit to the ongoing encouragement of others. Such factors would supposedly lead to an increase in self-confidence, and presumably be available again in similar situations of high salience.

The present study offers relatively unique findings in the attributional area through its investigation of two competing models in a real-life exam situation of high outcome salience, where consistency of attributions based on expected and actual outcome were a key element. Further research however should consider the dimension of controllability (Weiner, 1979, 1980). There appears to be increasing evidence that this dimension has an important influence on reactions to achievement-related outcomes (Forsyth and McMillan, 1981) particularly in terms of the development of learned helplessness among failure-prone students. It would also be useful for studies to investigate consistency of attributions over time within the ongoing achievement context of the classroom. There is a dearth of such information that is high in ecological validity, yet ultimately it would seem that verification of the causality of causal attributions requires such research.

ACKNOWLEDGMENTS.—This study was supported in part by funding from the Humanities and Social Sciences Research Fund, Massey University. The authors are grateful to Dr Beryl Hesketh for her assistance at various stages of the study and Dr Alison St. George for helpful comments on drafts of this paper.

REFERENCES

ARKIN, R. M., and MARUYAMA, G. M. (1979). Attribution, affect, and college exam performance. *J. educ. Psychol.*, 71, 85-93.

BAR-TAL, D. (1978). Attributional analysis of achievement-related behavior. *Rev. educ. Res.*, 48, 259-271.

BAR-TAL, D., and DAROM, E. (1979). Pupil's attributions of success and failure. *Child Dev.*, 50, 264-267.

BAR-TAL, D., and FRIEZE, I. (1977). Achievement motivation for males and females as a determinant of attributions for success and failure. *J. sex Roles*, 3, 301-313.

BAR-TAL, D., RAVIV, A., RAVIV, A., and BAR-TAL, Y. (1982). Consistency of pupils' attributions regarding success and failure. *J. educ. Psychol.*, 74, 104-110.

CHAPMAN, J. W., and BOERSMA, F. J. (1979). Learning disabilities, locus of control, and mother attitudes. *J. educ. Psychol.*, 71, 250-258.

COOPER, H. M., and BURGER, J. M. (1980). How teachers explain students' academic performance: a categorization of free response academic attributions. *Am. educ. Res. J.*, 17, 95-109.

DAVIS, M. H., and STEPHAN, W. G. (1980). Attributions for exam performance. *J. appl. soc. Psychol.*, 10, 235-248.

DEAUX, K., and FARRIS, F. (1977). Attributing causes for one's own performance: the effects of sex, norms, and outcome. *J. Res. Person.*, 11, 59-72.

DEPARTMENT OF EDUCATION (1982). *School Certificate Examination Statistics* 1981. Wellington: Department of Education.

DWECK, C. S., DAVIDSON, W., NELSON, S., and ENNA, B. (1978). Sex differences in learned helplessness II: The contingencies of evaluation feedback in the classroom, and III: An experimental analysis. *Dev. Psychol.*, 14, 268-276.

DWECK, C. S., and REPUCCI, N. D. (1973). Learned helplessness and reinforcement responsibility in children. *J. Person. soc. Psychol.*, 25, 109-116.

FEATHER, N. T. (1969). Attribution of responsibility and valence of success and failure in relation to initial confidence and task performance. *J. Person. soc. Psychol.*, 13, 129-144.

FEATHER, N. T., and SIMON, J. G. (1973). Fear of success and causal attributions for outcome. *J. Person.*, 41, 525-542.

FORSYTH, D. R., and MCMILLAN, J. H. (1981). Attributions, affect, and expectations: a test of Weiner's three-dimensional model. *J. educ. Psychol.*, 73, 393-403.

FRIEZE, I. H., and BAR-TAL, D. (1980). Developmental trends in cue utilization for attributional judgements. *J. appl. dev. Psychol.*, 1, 83-94.

FRIEZE, I. H., and WEINER, B. (1971). Cue utilization and attributional judgements for success and failure. *J. Person.*, 39, 591-605.

HAYS, W. L. (1981). *Statistics* (3rd edit.). New York: Holt, Rinehart and Winston.

KELLY, H. H., and MICHELA, J. L. (1980). Attribution theory and research. In ROSENZWEIG, M. R., and PORTER, L. W. (Eds), *Ann. Rev. Psychol.*, 31, 457-501.

LAWES, M. M. (1983). Causal attributions for outcome in School Certificate English. Unpublished Master's thesis, Massey University, New Zealand.

MCMAHAN, I. E. (1973). Relationships between causal attributions and expectancy of success. *J. Person. soc. Psychol.*, 28, 108-114.

MCMILLAN, J. H., and SPRATT, K. F. (1980). Causal attributions and affect in a real-life testing situation. Paper presented at the meeting of the American Educational Research Association, Boston, April 1980.

MEYER, J. P. (1980). Dimensions of causal attributions for success and failure: A multivariate investigation. In WEINER, B. (Ed.), *Human Motivation.* New York: Holt, Rinehart and Winston.

MILLER, D. T. (1976). Ego involvement and attributions for success and failure. *J. Person. soc. Psychol.*, 34, 901-906.

MILLER, D. T., and ROSS, M. (1975). Self-serving biases in the attribution of causality: fact or fiction? *Psychol. Bull.*, 82, 213-225.

NICHOLLS, J. G. (1975). Causal attributions and other achievement-related cognitions: effects of task outcome, attainment value, and sex. *J. Person. soc. Psychol.*, 31, 379-389.

OSTROVE, N. (1978). Expectations for success on effort-determined tasks as a function of incentive and performance feedback. *J. Person. soc. Psychol.*, 36, 909-916.

PEARL, R., BRYAN, T., and DONAHUE, M. (1980). Learning disabled children's attributions for success and failure. *Learn. Disability Q.*, 3, 3-9.

ROSENFIELD, D., and STEPHAN, W. G. (1978). Sex differences in attributions for sex-typed tasks. *J. Person.*, 46, 244-259.

SIMON, J. G., and FEATHER, N. T. (1973). Causal attributions for success and failure at university examinations. *J. educ. Psychol.*, 64, 46-56.

VALLE, V. A., and FRIEZE, I. H. (1976). Stability of causal attributions as a mediator in changing expectations for success. *J. Person. soc. Psychol.*, 33, 579-587.

WEINER, B. (1972). *Theories of Motivation: from Mechanism to Cognition.* Chicago: Markham.

WEINER, B. (1979). Theory of motivation for some classroom experiences. *J. educ. Psychol.*, 71, 3-25.

WEINER, B. (1980). The role of affect in rational (attributional) approaches to human motivation. *Educ. Researcher*, 9, 4-11.

(*Manuscript received 27th May,* 1983)

Br. J. educ. Psychol., 54, 189-205, 1984

INTERESTS, EFFORT, ACHIEVEMENT AND VOCATIONAL PREFERENCE

BY L. SJÖBERG
(Department of Psychology, University of Göteborg, Sweden)

SUMMARY. The components of interest in natural sciences and technology were studied in 100 students in the science and technology divisions of a Swedish secondary school. Also reported were perceived ability, success and invested effort in eight selected school subjects as well as vocational preferences and the reasons for these. Interests were found to be well accounted for by two factors—logical orientation and practical value. Interests and grades were strongly correlated but the correlation between interests and effort was more moderate. Structural analysis of the relationships between interest and achievement-related variables suggested that effort was mainly directed towards school subjects which students rated as important and in which they perceived a high degree of ability. Furthermore, effort led to success and success to interest. There was, however, little tendency for interest to lead to increased effort. Dispersion of interests correlated with a preference for natural science and technology as compared to the social sciences and the humanities. Interest in school subjects correlated at a moderate level with some of the vocational preferences; these were, however, little related to judged future job market prospects in vocations.

INTRODUCTION

INTEREST is a pleasant emotion which supports constructive activity (Tomkins, 1962). People often give interest as a reason for their vocational choice (Sjöberg, 1983). Despite the central role given to interest as a factor promoting action it has been relatively neglected in educational research. Most educational work on interest has considered it as an attitude variable and has neglected to analyse its components. Counselling researchers, on the other hand, have traditionally been concerned with the Strong and Kuder approaches (see e.g. Tyler, 1965) which are indirect and quite broad, not tied to specific study activity interests. Work on attribution of success in educational settings is relevant to the study of interest to some extent only, since it considers the student to be inquisitive mostly about reasons for outcomes rather than the contents of his or her studies (c.f. Weiner, 1979).

The possibility that interest plays a powerful role in guiding the selection and persistency of action calls for a more direct approach to the study of what interest is, how it develops and how it is sustained. Sjöberg (1983) studied the components of interest in eight high school subject areas, as well as attitudes to technology and vocational choice. It was found that grades and interests in specific school subjects correlated positively, except in the case of very high achievers for whom the correlation was negative.

In attribution theory (Kelley, 1967) it is held that the attribution of achievement is due, among other things, both to effort and ability. Since the earlier study (Sjöberg, 1983) showed moderately strong relations between choice and interest and achievement, it was considered desirable to include a more differentiated view on the attribution of achievement. Specifically, the grade-interest correlation found in the earlier study may have been mediated by perceived effort and ability. Grades may contribute more to interest on high levels of perceived ability rather than effort because ability is typically a more permanent and important characteristic of the self than the effort exerted in any particular task. In other words, grades can be assumed to be construed by the subjects as more diagnostic if achievement is attributed to ability rather than effort.

Moderating effects of ability and effort on the grade-interest relationship can be investigated if data on perceived effort and ability are available. If relationships among variables are, on the other hand, linear with no interactions, it is appropriate to attempt another type of structural analysis—applying, e.g., Jöreskog's LISREL procedure (Jöreskog, 1978; Jöreskog and Sörbom, 1981).

Interest could not predict vocational choice in the earlier study, in spite of a large number of students stating that " interest " was their main reason for choice. The reason may partly have been that the response category " Don't know " was allowed and it was frequently chosen, indicating considerable uncertainty of choice and resulting in incomplete data. It was also striking that vocational choice was little related to rated job market prospects. More data on job prospects seemed to be called for.

The purpose of the present study was to investigate further the relationship among interests, grades, and choice, taking into account the need for an improved choice scale, more data on job prospects, and a finer differentiation of the attribution of achievement. In particular, this study investigates the structural relations among the variables with special reference to how interest arises and if it influences effort and performance.

METHOD

Subjects

The subjects were 100 high school (in a Swedish " gymnasium ") students; 71 boys and 29 girls. They were students in natural science (56) or technology (44) divisions. Their mean age was 18 years. They participated in the study at the end of their last school year. They were enrolled in three different high schools which recruited pupils mostly from middle and upper middle classes in the city of Gothenburg and its suburbs.

Questionnaire

The questionnaire contained questions about background data, grades in selected school subjects, parents' occupations, leisure interests and some aspects of the planning of studies. It included 14 attitude questions concerned with society and technology, defining two factors found by Sjöberg (1983). The two factors may be termed " limited perspective of technologists " and " contributions of technology to society ". Then followed a section asking for category ratings of eight subject areas, viz. history, mathematics, social science, physics, Swedish, biology, English and technology. The school subjects were rated in the following aspects: general interest, personal importance of the subject, importance for society, experienced success, importance of success in the subject, value in leading to insight, facilitation of own activity, practical value, surprising insights, attraction of logic and structure, remaining unsolved problems, effort invested, and perceived ability to succeed in the subject. Except the last two, these aspects were also used by Sjöberg (1983) and they were originally formulated to cover aspects found in a preliminary interview study which attempted to find the components creating and sustaining interest in mathematics and science.

The final section of the questionnaire called for data on vocational choice. The students rated, on a three-category scale with the categories " yes ", " maybe " and " no ", five varieties of engineering (electronics, engineering physics, chemistry, road construction and mechanical engineering), physician, dentist, physiotherapist and teacher of mathematics and science. They then stated which was the most important factor in determining their choice, who was most important in promoting their interest in school subjects, who affected their vocational choice and whether they felt their interests were similar to those of their parents. Finally, the students rated the future prospects of getting a stimulating and well-paid job after having finished their studies to become a graduate engineer, physician, dentist, physiotherapist or teacher of mathematics and science.

Data were, with few exceptions, complete. (Students in the science division did not rate technology since that was not a subject in their curriculum.) Variables were treated as ordered except for those referring to occupation, leisure interest, reasons for choice and who had affected choice and interests. Data analysis was carried out in a quantitative mode with few exceptions.

Attitude factors were scored as in Sjöberg (1983) and so were interest ratings. The latter were grouped as General Interest, Logic and Practical Value, based on factor analyses reported in the earlier study.

RESULTS

Grades

There were high positive correlations among grades in the five school subjects. Those who reported studying and planning more also had higher grades; those who delayed exams had lower. Grades correlated with choice of more prestigious fields of future study; the choice of mechanical engineering correlated negatively with grades. The latter finding may be due to variation between the two high school divisions. Pupils in the division of technology had lower grades and tended to favour the mechanical engineering division of the engineering schools (cf. Sjöberg, 1983). Those who had better grades were also more optimistic about their future prospects.

Grades correlated with interest factors as given in Table 1. Note that the practical value of the school subject correlated only slightly with grades while the major interest dimensions of General Interest and Logic (for mathematics and physics) correlated more, as did perceived effort and perceived ability.

In the previous study (Sjöberg, 1983) a tendency towards an inverse U-shaped relationship between interest and mathematics grades was found. To test the generality of this finding one-way analyses of variance were performed with the three interest factor scores as dependent variables and grades as independent variables. (Interest factors were first averaged across all eight school subjects.) There were no tendencies towards deviations from linear trends in these data; higher grades tended to imply higher interest scores.

A further possibility suggested in the introduction was that of moderating effects of perceived ability and effort on the grade-interest relationship. To test for this

TABLE 1

CORRELATIONS BETWEEN INTEREST AND SPECIFIC VARIABLES AND GRADES

				Specific variables		
	General interest	Practical value	Logic	Success	Effort	Ability
Mathematics	44**	13	48**	68**	43**	47**
Physics	49**	20	45**	67**	45**	33**
Social science	26**	15	13	57**	47**	31**
English	30**	21*	22*	56**	30**	41**

Decimal point omitted. * $P<0.05$ ** $P<0.01$

hypothesis, regression analyses with grades as the dependent variables were performed on the four school subjects for which data were available on grades, interests, effort and ability (mathematics, physics, social science and English). The analyses were done in three steps where ability and effort were entered first, then interest factor scores and finally the ability-effort product to check for the interaction hypothesis. In no case, however, did the product term contribute significant variance not already explained by the linear regressions. Hence, there was no indication of a moderator effect. Further analysis of relations among grades, interest, ability and effort is provided below in the section on structural modelling.

Work habits

As already noted, work habits correlated positively with grades. Those with orderly habits and more time invested in school work were also more optimistic about the future. Work habits correlated with perceived effort and systematically, but not strongly, with interest factors.

Attitudes

Attitudes to technology did not correlate with grades in mathematics and physics. Engineering choice correlated positively with attitudes. Choice of health professions and teaching tended to be combined with negative attitudes to technology. Those who evaluated technology as beneficial were also more optimistic about future job prospects.

There were consistent correlations between interest factors and attitudes towards mathematics, physics, technology and biology (Table 2). As can be seen in the table, interest in mathematics, physics and technology was combined with a positive attitude to technology, while the reverse was true for interest in biology.

TABLE 2

CORRELATIONS BETWEEN INTEREST AND ATTITUDE DIMENSIONS

	Factor 1			Factor 2		
	General interest	Practical value	Logic	General interest	Practical value	Logic
Mathematics	11	19	10	−25*	−10	−09
Physics	25*	29*	24*	−30*	−27**	−17
Technology	43**	32**	38**	−20	−26*	−24
Biology	−31**	−06	−06	16	−02	−02

Decimal point omitted. * $P < 0{\cdot}05$ ** $P < 0{\cdot}01$

Note: 1. Contributions of technology to society, positive.
2. Limited perspective of technologists, negative.

Interest

Subjects rated technology, physics and mathematics high in General Interest while history, social science and Swedish were low. English and biology occupied intermediary positions. In particular, history was rated low both in Practical Value and Logic. To test the differences, physics and mathematics interests were pooled and so were history, social science, Swedish and English. The two pooled interest scores were termed Math and Hum. The mean scores are given in Table 3 as well as the results of t-testing selected differences. All three tests gave highly significant differences.

TABLE 3

MEAN POOLED MATHEMATICS AND PHYSICS (MATH) AND HUMANITIES (HUM) SCORES

	Math	Hum	t¹	Significance of difference, two-tailed
General interest	5·53	4·72	6·93	<0·001
Practical value	5·54	4·93	5·89	<0·001
Logic	4·23	3·32	7·14	<0·001

Note: Df = 99.

Differences between Math and Hum were computed and correlated with other variables. They correlated positively with grades, effort, success and ability in mathematics and physics and with technically oriented hobbies. Relations with vocational choice will be mentioned in a later section.

There were only small differences among school subjects in average perceived success and perceived ability. Perceived effort was highest for technology, physics, mathematics and biology, and lowest for history.

General Interest could be well accounted for by Practical Value and Logic. Multiple correlations and β weights are given in Table 4. The multiple Rs are all around 0·70. Taking measurement errors into account—Sjöberg (1983) reported reliabilities around 0·85—this means that less than 10 per cent of the variance of General Interest is left when Practical Value and Logic are held constant. It is also striking in Table 4 that Logic dominates interest in the natural sciences and technology while Practical Value is a more important aspect for languages and social science. It is especially noteworthy that this trend holds also for technology. The results agree with the interviews in the pilot study in which practical value was given a rather low weight in accounting for interest in natural sciences and technology.

It may be added that standard deviations of interest factor scores varied little between school subjects (with the exception that mathematics had a larger standard deviation in all three dimensions) and that β weights did not reflect average level; e.g., technology had the highest average rating in Practical Value and a low β weight.

TABLE 4

MULTIPLE CORRELATIONS AND β WEIGHTS FOR PREDICTING GENERAL INTEREST FROM PRACTICAL VALUE AND LOGIC FACTOR SCORES

	β weights		
Subject	Practical value	Logic	Multiple correlation
Mathematics	33	46	67
Physics	29	52	70
Technology	29	51	72
Biology	27	48	67
History	40	42	76
Social science	44	34	69
Swedish	55	29	73
English	45	40	69

Decimal point omitted

Intercorrelations among subjects in General Interest are given in Table 5. Clearly, mathematics, physics and technology tended to form a cluster, as did history, social science, Swedish and English. Biology was present in both clusters while technology interest correlated *negatively* with interest in the other subjects except mathematics and physics. When it comes to perceived effort and ability, the picture is different (Table 6). Here, all subjects, including technology, were positively correlated. Apparently pupils saw themselves as generally able and involved, though they clearly distinguished among subjects in terms of interest.

Relations between Perceived Effort and interest factors are given in Table 7. It is noteworthy that General Interest correlated most strongly with effort while Logic was pronounced as a secondary factor for mathematics and physics and Practical Value for technology and English. Note that the total effort, as reflected in the number of hours devoted to homework, did not correlate as clearly with interest as did subject area perceived effort.

TABLE 5

CORRELATIONS AMONG SCHOOL SUBJECTS, GENERAL INTEREST

	1	2	3	4	5	6	7	8
1. Mathematics	—							
2. Physics	72	—						
3. Technology	30	31	—					
4. Biology	30	24	−23	—				
5. History	08	04	−21	43	—			
6. Social science	13	10	−18	26	65	—		
7. Swedish	26	18	−17	39	57	50	—	
8. English	11	15	−04	28	33	22	36	—

Decimal point omitted

TABLE 6

INTERCORRELATIONS AMONG SUBJECTS. BELOW DIAGONAL: PERCEIVED ABILITY. ABOVE DIAGONAL: PERCEIVED EFFORT

	1	2	3	4	5	6	7	8
1. Mathematics	—	76	29	59	44	46	52	57
2. Physics	76	—	23	55	28	35	42	39
3. Technology	48	54	—	04	06	19	17	28
4. Biology	56	58	38	—	58	44	48	44
5. History	25	23	32	43	—	69	48	34
6. Social science	28	27	07	26	70	—	57	24
7. Swedish	46	39	22	50	70	55	—	45
8. English	36	35	34	56	39	27	52	—

Decimal point omitted

TABLE 7

CORRELATIONS BETWEEN INTEREST FACTORS AND EFFORT

	General interest	Practical value	Logic
Mathematics	40**	21*	33**
Physics	33**	15	34**
Technology	53**	42*	38*
Biology	57**	26*	48**
History	39**	23*	24*
Social science	41**	18	22*
Swedish	40**	30*	28*
English	50**	35**	27*

Decimal point omitted. * P<0·05 ** P<0·01

Three-mode component analysis

A three-mode component analysis was carried out in order to study more closely the structuring of the variables and the subject areas. The modes were variables, subject areas (excluding technology since that area was not studied by all individuals) and individuals. The TUCKALS 3 algorithm of Kroonenberg (Kroonenberg and DeLeeuw, 1980) was used to analyse the data. Three components were extracted for the variables, two for the subject areas and five for the individuals, after standardising the variables to equal means and variances and subject areas to equal means. The proportion of the total sum of squares accounted for by these components was, for the variables, individuals and subject areas, 0·62, 0·50 and 0·64, respectively.

For the variables, the third component was small and could not be interpreted. The plot of the first two components can be seen in Figure 1. The variables cluster according to the factors found in the earlier study with the exception of those measuring practical value. (As noted above, their status in the earlier factor analyses was also uncertain.) The variables effort and perceived capacity tended to join the cluster of the general interest variables. With the exception of the variables measuring practical value the three-mode analysis supported the classification of variables in factors scores as utilised here.

The subject areas are plotted in a two-dimensional component space in Figure 2. The transition from languages and humanities to natural sciences is quite clear, supporting the results presented above.

FIGURE 1

THE FIRST TWO COMPONENTS FOR THE VARIABLE MODE IN THE THREE MODE FACTOR ANALYSIS

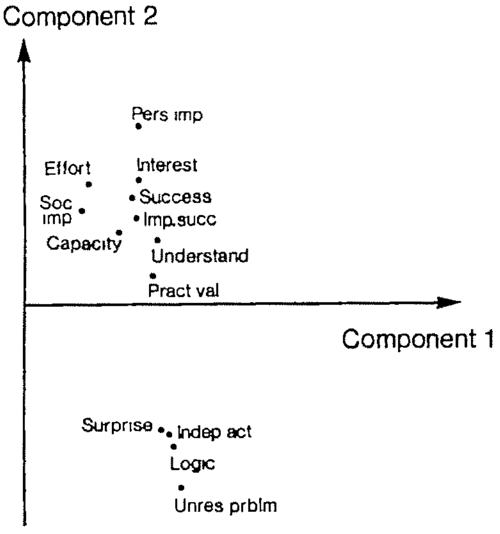
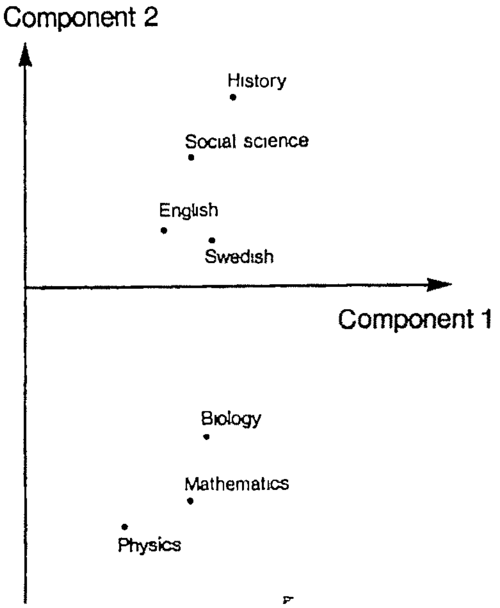


FIGURE 2

THE TWO COMPONENTS FOR THE SUBJECT AREA MODE IN THE THREE MODE FACTOR ANALYSIS



The abbreviations of the names of the variables in Figure 1 and their corresponding more complete descriptions are as follows: Interest (general interest), Pers. imp. (personal importance of the subject), Soc. imp. (importance for society), Success (experienced success), Imp. succ. (importance of success in the subject), Understand (value in leading to insight), Indep. act. (facilitation of own activity), Pract. val. (practical value), Surprise (surprising insights), Logic (attraction of logic and structure), Unres. prblm. (remaining unsolved problems), Effort (effort invested), Capacity (perceived ability to succeed in the subject).

Models of grade-interest relations

The relations between grades and pertinent experience variables will now be modelled more closely. The following variables will be used:

(1) *Grade.*
(2) *Interest* rating.
(3) *Effort* ratings. Effort ratings were analysed both in raw form and after adjustment to a common mean for all subjects (ipsatised form). The latter transformation was applied because this author was interested in how subjects distributed effort apart from their total level of effort.
(4) *Ability* rating.
(5) *Perceived success* ratings.
(6) *Importance*, obtained from the ratings of how important success in the subject area was to the student.

These variables were available for four school subjects, viz. mathematics, physics, English and social science. A separate analysis was therefore performed on each of these subject areas.

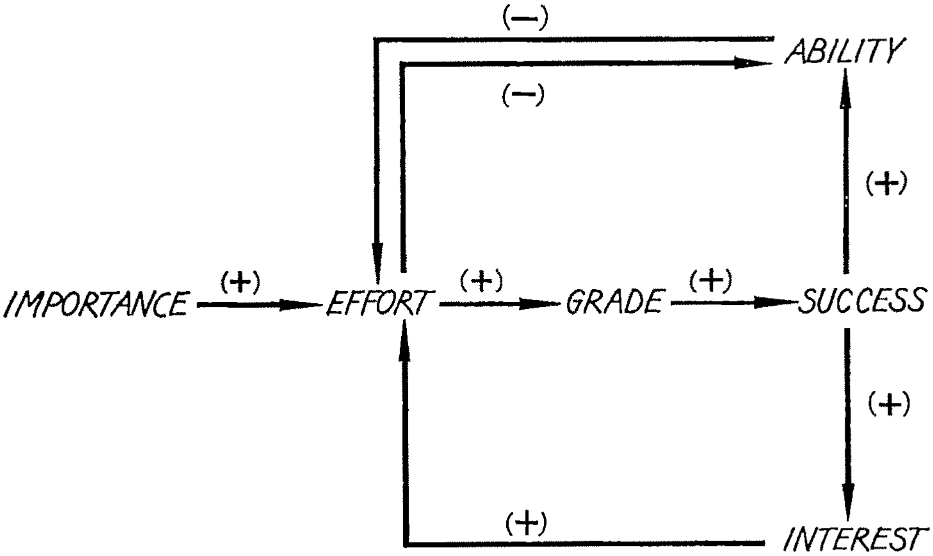
The first model to be analysed is given in Figure 3. The model assumes a common-sense theory of action (Heider, 1958) for the effort-grade-success-ability relations but it adds interest (not included in Heider's analysis of action) and importance of the task.

Causal flows are postulated for 7 out of 15 pairs of variables; they will be commented upon in the following:

(1) Effort affects grade in a positive manner.
(2) Grade affects perceived success.
(3) Success leads to interest.
(4) Success also leads to perceived ability. Effort invested affects perceived ability.
(5) Interest leads to effort.
(6) Ability, together with the goal implicit in the rated importance of the subject, affects effort.

FIGURE 3

SUGGESTED MODEL FOR RELATIONS AMONG ACHIEVEMENT AND INTEREST VARIABLES



Note that the effort-ability relation is assumed to be two-way. The notion of a distinct goal set by the student calls for a negative relation between effort and ability. Those who perceive themselves to be very able should need to invest less effort and effort invested should be an indication of ability; those who see themselves as working more should account for this by attributing less ability to themselves.

The model was analysed with Jöreskog's LISREL V program which allows for the testing of structural models on correlation or covariance matrices (Jöreskog and Sörbom, 1981). The LISREL procedure provides maximum likelihood estimates of free parameters and a chi^2-test of goodness of fit of the postulated model. The model was applied to the four school subjects in turn. All analyses were performed both on raw effort ratings and on such ratings after adjustment to a common mean (ipsatised) in order to distinguish more clearly the allocation of effort. It was found that the lack of fit was highly significant in all cases. Therefore, an attempt was made to find new models that might fit better to data by stepwise introduction of new parameters where bad fits could be located. The results of this procedure were unsatisfactory for three of the four subject areas.

However, LISREL V allows for a different approach. Some relations may be specified as prohibited (not allowed to be estimated freely) and the others estimated through an iterative procedure, in which each new stage involves estimation of that pairwise relation in the previous stage which fitted maximally badly. In this way, convergence to well-fitting models was obtained. The prohibited relations are given in Figure 4, the obtained models in Figures 5-8. The figures give results for raw effort ratings except for English. In the latter case results for ipsatised effort are given since they were clearer and more in line with the other three analyses.

FIGURE 4

PROHIBITED RELATIONSHIPS FOR ITERATIVE LISREL MODEL FITTING

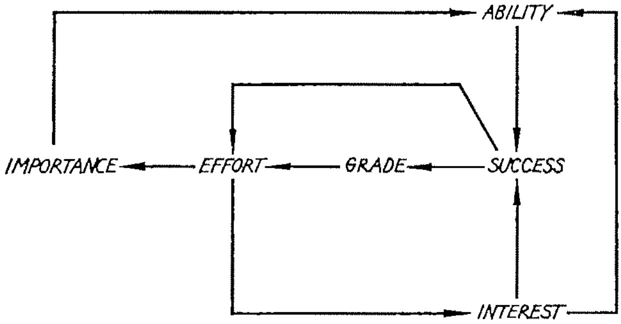


FIGURE 5

STRUCTURAL MODEL FITTED FOR MATHEMATICS

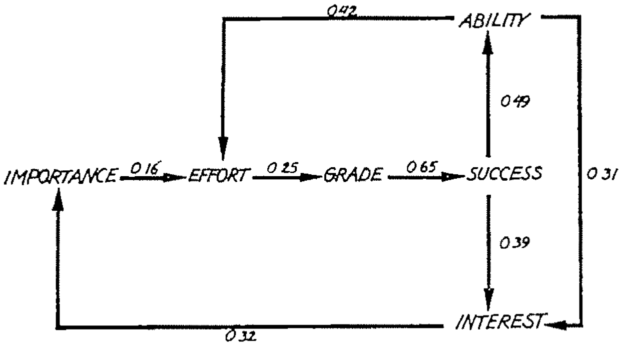


FIGURE 6

Structural Model Fitted for Physics

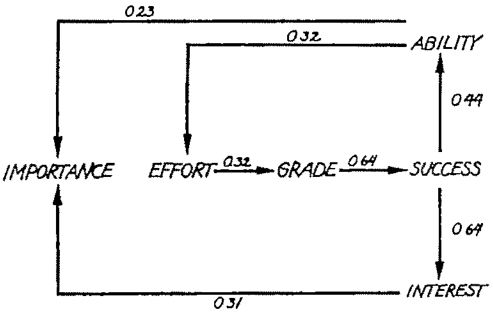


FIGURE 7

Structural Model Fitted for Social Science

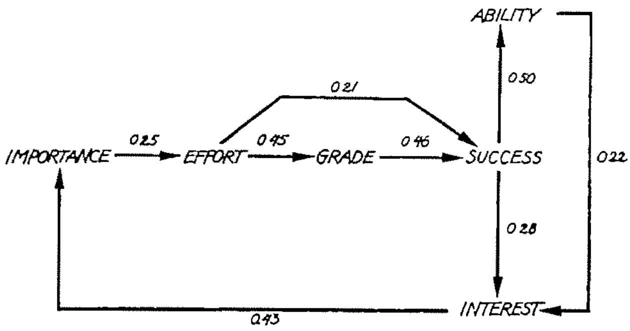
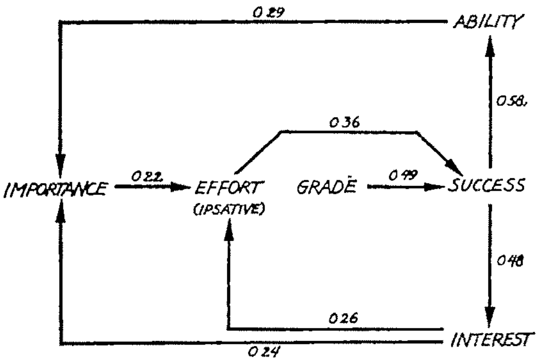


FIGURE 8

Structural Model Fitted for English



Comparing the models in Figures 5-8 with the originally postulated one (see Figure 3) it is found that there are some interesting deviations in data that were not anticipated when model 1 was formulated. The following points may be noted:

(1) There are only weak traces of direct feedback from interest to effort. Some indirect feedback occurs and is mediated by the importance ratings.

(2) In no case could a negative feedback from ability to effort be established. In fact, the relationship between these two variables was asymmetric, implying increased effort with increased perceived ability, in two of the cases (mathematics and physics).

(3) Perceived ability seemed to affect interest positively in the cases of mathematics and social science.

(4) There were some differences between the school subjects. In particular the effort-ability relation was established mainly for mathematics, physics and English and the feedback to importance was most clear for mathematics and social science.

(5) For three subject areas there was a direct relation from effort to perceived success which did not " pass through " the grade variable.

(6) There were no clear and systematic differences between raw and ipsatised effort ratings.

Summing up, the analyses suggest that success and perceived ability lead to increased interest but that interest is less strongly, if at all, likely to lead to increased effort. Furthermore, effort seems not to be increased due to a feeling of less ability; on the contrary effort, if at all related to ability, is related in the *opposite* manner to the one postulated: effort is increased as a function of perceived ability.

Reasons for interest and choice

The following question was asked: Which is the most important factor in deciding your vocational choice? The overwhelming majority, 83 per cent, answered " Interest ". Other scattered answers were " job market " (1 pupil), " salary " (1 pupil) and various combinations of factors. Next, pupils were asked who was most important in producing interest in school work. The dominating answer here was " teacher ", 37 per cent. Parents and siblings were also rather important: 21 per cent pointed to them as the major agent. However, when it was asked who influenced choice of vocation, " teacher " was given as the answer in only 10 per cent of the cases, family members in 40 per cent. Peers were stated as the major agent in influencing vocational choice by 22 per cent of the pupils. In 21 per cent of the cases, pupils said nobody was a major agent influencing them, " they made the choice by themselves ".

Thus, teachers were seen as influencing interest, while family and friends influenced vocational choice, which was however perceived by many as internally rather than externally determined.

Subjects were asked if their interests were similar to those of their parents. Many, 50 per cent, denied any similarity; most of the others, 28 per cent of the whole group, said " both ".

Choice of vocation and future job market prospects

Choice of future vocation correlated to some extent with interest factors. Results for mathematics and biology are given in Tables 8 and 9. Partial correlations between choice with interest, grades constant, are also given in these tables.

TABLE 8

CORRELATIONS BETWEEN CHOICE OF ENGINEERING AND MEDICINE AND MATHEMATICS INTEREST, SUCCESS AND PERCEIVED ABILITY. VALUES WITHIN PARENTHESES ARE PARTIAL CORRELATIONS WITH GRADES CONSTANT

	Ratings of mathematics				
Choice	General interest	Practical value	Logic	Success	Ability
Graduate engineer electronics	39** (27*)	28* (24*)	40** (28*)	29** (05)	09 (−06)
Graduate engineer engineering physics	44** (30**)	23* (18)	41** (27*)	36** (09)	04 (−15)
Graduate engineer chemistry	29** (12)	01 (−07)	28* (11)	32** (03)	−01 (−21)
Graduate engineer road construction	14 (08)	−02 (−05)	06 (−01)	15 (07)	15 (−23*)
Graduate engineer mechanical engineering	02 (09)	−10 (−08)	−02 (08)	03 (09)	−31** (−29**)
Physician	04 (−16)	−07 (−14)	19 (−02)	33** (10)	28* (16)

Decimal point omitted. * P<0·05 ** P<0·01

TABLE 9

CORRELATIONS BETWEEN CHOICE OF ENGINEERING AND MEDICINE AND BIOLOGY INTEREST, SUCCESS AND PERCEIVED ABILITY. VALUES WITHIN PARENTHESES ARE PARTIAL CORRELATIONS WITH GRADES CONSTANT

	Ratings of biology				
Choice	General interest	Practical value	Logic	Success	Ability
Graduate engineer electronics	17 (01)	19 (14)	18 (06)	12 (−07)	04 (−12)
Graduate engineer engineering physics	14 (−06)	05 (−02)	15 (00)	20 (−01)	16 (−27*)
Graduate engineer chemistry	24* (06)	02 (−06)	22 (08)	30** (12)	02 (−05)
Graduate engineer road construction	13 (07)	09 (07)	10 (06)	19 (14)	−04 (−07)
Graduate engineer mechanical engineering	−16 (−11)	−12 (−10)	−04 (01)	−02 (06)	−28* (−26*)
Physician	28* (14)	22 (17)	26* (15)	38** (24*)	32** (27*)

Decimal point omitted. * P<0·05 ** P<0·01

It is seen in Table 8 that interest in mathematics contributes to choice of two of the more prestigious engineering divisions, also with grades constant. For biology, the same is true for choosing to become a physician. Other correlations between choice and interest can be explained by variation in grades.

A somewhat more clear-cut picture is afforded by Table 10 which gives correlations between Math-Hum differences in the three interest factors and choice. It is clear for four out of five engineering orientations that the Math-Hum difference has some predictive power; this was not true for any of the other vocational options studied here. It may be noted, in addition, that interest dispersions correlated less

TABLE 10

CORRELATIONS BETWEEN VOCATIONAL PREFERENCE AND MATH-HUM DIFFERENCES IN THE THREE INTEREST FACTORS

	Math-Hum difference		
Choice	General interest	Practical value	Logic
Graduate engineer electronics	23*	09	28*
Graduate engineer engineering physics	30**	09	40**
Graduate engineer chemistry	15	04	26*
Graduate engineer road construction	11	−04	03
Graduate engineer mechanical engineering	30**	22**	07
Physician	−05	−07	03
Dentist	−12	−08	−13
Physiotherapist	−07	00	00
Teacher	15	−06	14

Decimal point omitted. * P<0·05 **P<0·01 with vocational preferences than Math-Hum differences, in spite of sizeable correlations between these two sets of discrimination measures.

Future job market prospects were rated for five types of vocation. They intercorrelated rather strongly, giving some credence to a factor of general optimism (cf. Sjöberg and Biel, 1983). There were virtually no correlations between job market prospects and vocational choice, a surprising finding. Vocational choice seemed to be more related to interest—a finding agreeing with the fact that pupils themselves said that interest was the major factor in their vocational choice.

Sex differences

There were some suggestive sex differences. Girls reported that they worked harder and more systematically. They chose health professions more often than boys and were more optimistic about the future and more negative to technology. Correlations between sex and interest in mathematics, physics and biology are given in Table 11. It is clear that girls were more interested in biology, boys in physics and, possibly, mathematics. It is noteworthy that girls reported more effort in all subjects, in spite of lower interest in mathematics and physics. They also perceived less success in physics.

A final note on sex differences is due to the question in which pupils were asked about whose interest they found their own interest to be most similar to. The results are given in Table 12. The differences between sexes is significant (chi^2 = 11·64, df = 3, $P<0·05$). Boys perceived themselves as more like their fathers, girls as more like their mothers. More interesting, however, is the finding that boys more often asserted their independence by denying any similarity with their parents.

TABLE 11

CORRELATIONS BETWEEN SEX AND INTEREST AND THREE SPECIFIC VARIABLES

	General interest	Practical value	Logic	Success	Effort	Ability
Mathematics	19	13	07	07	−24*	−04
Physics	27**	20*	14	28**	−09	15
Biology	−45**	−19	−24*	−19	−33**	−25*

Note: Male was scored as 2, female as 1 so a positive correlation means that males had a higher mean than females.
Decimal point omitted.

TABLE 12

NUMBER OF ANSWERS TO THE QUESTION
" WHOSE INTERESTS ARE SIMILAR TO YOUR OWN? "

Response	Girls	Boys
Father	1	13
Mother	4	1
Both parents	11	17
None	12	38

Hobbies and parent's occupation

Occupations of parents and hobbies were scored dichotomously as to whether they were concerned with science and technology or not. The variables thus obtained were correlated with other data. There were no sizeable correlations with parent's occupation. As for hobbies, it was striking that interest level was *negatively* related to scientific hobbies, while interest dispersion was positively related: see Table 13 which also gives correlations with Math-Hum differences. The former finding is hard to explain; the latter points, once more, to scientific interest being a simultaneous espousal of sciences and a rejection of humanities.

TABLE 13

CORRELATIONS BETWEEN HOBBIES, MEAN INTEREST SCORES, DISPERSIONS OF INTEREST SCORES AND MATH-HUM DIFFERENCES IN THE THREE INTEREST FACTORS

Interest factor	Mean	Dispersion	Math-Hum
General interest	−29**	46**	40**
Practical value	−23*	50**	34**
Logic	−21*	30*	32**

* $P<0{\cdot}05$ ** $P<0{\cdot}01$

Note: Scored as 1 for nonscientific and 2 for scientific-technological.

DISCUSSION

Interest could in this study be well accounted for by the factors Logic and Practical Value. It was striking that interest in natural sciences and technology was so clearly dominated by the Logic factor while interest in languages and social sciences was mostly determined by Practical Value. Also, scientific and technological interests tended to imply a rejection of the humanities.

The present findings on grades and interest agree well with those in the earlier study (Sjöberg, 1983). Interests and grades were rather strongly correlated, and interest led, according to the structural models, directly or indirectly to increased effort. That is a finding inconsistent with work on the effects of external reward on intrinsic motivation (Deci and Ryan, 1980). Work on that topic is reviewed elsewhere (Sjöberg, 1982); let it suffice here to make a few statements.

Bates (1979) argued that the effects of reward on school motivation probably are quite complex and cannot be safely generalised from the type of experimental work reviewed by Deci and Ryan (1980). The direct attempts at educational field studies of effects of reward systems can be criticised on methodological grounds. Deci, Nezlek and Sheinman (1981) who studied effects on teachers' attitudes to reward on pupils' intrinsic motivation failed to get predicted changes in intrinsic motivation. The relation they found could be due to variation in norms across schools. A follow-up gave some further support (Deci *et al.*, 1981) but the main conclusions were based on variation between two teachers only.

The effect of decreased interest due to extrinsic reward seems to be critically dependent on subjects' belief that the reward behaviour was merely instrumental. The work in the area has, however, not given very clear-cut results and most of the theories put forward are unclear and not well supported empirically (Sjöberg, 1982).

A positive grade-interest relationship has two sides: those who achieve well develop interest while those who achieve poorly tend to lose it. The presence of punishment in " reward " systems seems to be quite important for understanding any detrimental effects of such systems on interest. The punishment aspect can be more or less salient, of course. Relative grades seem, to take an example of some current interest in Sweden, to be more punitive than grades based on absolute criteria of achievement. It is likely that punitive elements of reward systems are more important for the erosion of interest than perceived instrumentality of externally rewarded behaviour.

Since the relationship between perceived success and interest was linear there are, hence, two sides to the grade—interest issue: success breeds interest and failure breeds lack of interest. Abolishing grades could make performance feedback less salient and thus have the effect of decreasing variation in interest. A more constructive policy would be to enhance the performance of low achievers, thereby contributing to their learning, creating positive experience and subsequent interest. This may be easier to say than to do; one interesting example is, however, afforded by the use of planning illustrated by a study of low achievers in elementary mathematics by Bandura and Schunk (1981).

Many studies have shown a positive relation between interest and reward (Reiss and Sushinski, 1975; Bandura, 1982). Such an effect is likely especially when reward is conceived by the rewardee as a valid sign of some important type of competence. The structural models fitted to the present data also showed a tendency for perceived ability to be positively related to interest. The net effect of a reward thus depends on how salient it is as a signal of competence.

McGraw (1978) reviewed evidence for a detrimental effect of rewards on performance in non-routine attractive tasks. Within the same task, there may also be a variation between individuals; evaluation apprehension may lead to a superficial learning set (Fransson, 1977).

The problem with the evidence cited by McGraw and his interpretation of it is that the subjects' decision-making is neglected. Perhaps people *choose* to opt for many quickly produced answers in order to obtain rewards; this may or may not lead to decreased performance depending both on scoring rules and on the veridicality of the subjects' perceptions of their success probabilities and on their utilities. They may, e.g., simply like the rewards more than the task activity. These factors were not discussed by McGraw. To this author's knowledge, no study has demonstrated that rewards make subjects perform worse in achieving what they attempt to achieve. A study of this problem must collect data on intentions as well as performance.

The finding of a curvilinear relationship between mathematics grade and interest reported by Sjöberg (1983) was not replicated here; all grade-interest relations appeared to be linear. Furthermore, ability-effort combinations did not moderate the grade-interest relation. The three-mode component analysis gave results in accordance with the more straightforward analyses at the level of the manifest variables. The structural analysis gave some unexpected results in that students showed no tendency to trade off perceived ability and invested effort. A methodological improvement is called for, however, since it is not quite clear whether the ratings of effort were absolute and did not require the subjects to take into account the distribution of a limited amount of time and mental resources. The attempt to get ipsative effort data may not have been sufficient to counteract the effects of the way rating instructions were formulated.

The term "effort" may carry too many negative connotations to catch a possible relation between interest and the amount of time and energy invested in a subject, but this suspicion is partly contradicted by the positive relation between effort and perceived ability.

The structural analysis suggests a mode of functioning which is little related to strategic decisions as to how to distribute effort and develop interest. It seems that students tended to work more if they perceived themselves as more able, such work leading to success and interest. Part of the perceived success was not mediated by grades in the case of social science; it may be that students did not feel that grades really reflected their true performance.

The lack of tactical considerations inferred from the models may be due to the fact that the Swedish system of admittance to post-secondary education does not encourage pupils to choose the demanding natural science and technology divisions of the high school, nor does it encourage specialisation. The present sample thus consisted of students who had demonstrated a lack of tactical considerations in their choice of high school division. Further study with pupils of the short, and less demanding, two-year divisions of the high school would therefore be of interest. Also, it would be of interest to study pupils from a wider range of social backgrounds.

We also do not know if the high school grades measure superficial or depth learning. The ability to learn may, of course, be quite different depending on the contents of the subject and the grading procedures. Also, the definition of success may vary among pupils.

Vocational choice was more predictable from interests in the present study than in the previous one. It is remarkable that future job market prospects still had *no* correlation with vocational choice which was, by an overwhelming majority, said to be based on " interest ". Perhaps students, just as drivers of automobiles (Svenson, 1981), believe that their own future is more positive than that of people in general. It is also striking that the students, especially the boys, so often stated their independence when it came to choice and interest.

The relatively low importance attributed to teachers in influencing vocational choice in the present data agrees with results by Dupont *et al.* (1979). The latter authors did not, however, allow for the subjects to attribute their choice to the own person, a response that dominated the present data. The importance of the parents, especially the mother, is stressed by Dupont *et al.* (1979) and it agrees with data on political attitudes (e.g., Oskamp, 1977).

It was found in the present as well as in the previous study that dispersion of interest was related to a preference for technology and science. It seemed that such a preference tended to go together with a negative view of social science and the humanities. It would be of interest to see if this holds also in a non-selected group of students.

ACKNOWLEDGMENTS.—This study was supported by a grant from the Office of the Chancellor of the Swedish Universities. The author is grateful to Sten-Olof Brenner and Dag Sörbom for helping in the LISREL analyses and to Pieter Kroonenberg for the three-mode factor analysis. The paper was written while the author was associated with the Faculté Universitaire Catholique de Mons (FUCAM), Belgium. It was also supported by the European Institute for Advanced Studies in Management.

REFERENCES

BANDURA, A. (1982). Self-efficacy mechanism in human agency. *Am. Psychol.*, 37, 122-147.

BANDURA, A., and SCHUNK, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *J. Person. soc. Psychol.*, 41, 586-598.

BATES, J. A. (1979). Extrinsic reward and intrinsic motivation: a review with implications for the classroom. *Rev. educ. Res.*, 49, 555-576.

DECI, E. L., NEZLEK, J., and SHEINMAN, L. (1981). Characteristics of the rewarder and intrinsic motivation of the rewardee. *J. Person. soc. Psychol.*, 40, 1-10.

DECI, E. L., and RYAN, R. M. (1980). The empirical exploration of intrinsic motivational processes. In BERKOWITZ, L. (Ed.), *Advances in Experimental Social Psychology* (Vol. 13). New York: Academic Press.

DECI, E. L., SCHWARTZ, A. J., SHEINMAN, L., and RYAN, R. M. (1981). An instrument to assess adults' orientations towards control versus autonomy with children: reflections on intrinsic motivation and perceived competence. *J. educ. Psychol.*, 73, 642-650.

DUPONT, J.-B., GENDRE, F., BERTHOUD, W., and DESCOMBES, J.-P. (1979). *La Psychologie des Intérêts*. Paris: Presses Universitaire de France.

FRANSSON, A. (1977). On qualitative differences in learning: IV. Effects of intrinsic motivation and extrinsic test anxiety on process and outcome. *Br. J. educ. Psychol.*, 47, 244-257.

HEIDER, F. (1958). *The Psychology of Interpersonal Relations*. New York: Wiley.

JÖRESKOG, K. G. (1978). Structural analysis of covariance and correlation matrices. *Psychometrika*, 43, 443-477.

JÖRESKOG, K. G., and SÖRBOM, D. (1981). LISREL V—Analysis of linear structural relationships by maximum likelihood and least squares methods. *Research Report* 81-8, Department of Statistics, University of Uppsala.

KELLEY, H. H. (1967). Attribution theory in social psychology. In LEVINE, D. (Ed.), *Nebraska Symposium on Motivation*. Lincoln: University of Nebraska Press.

KROONENBERG, P. M., and DELEEUW, J. (1980). Principal component analysis of three-mode data by means of alternating least squares algorithms. *Psychometrika*, 45, 69-97.

MCGRAW, K. O. (1978). The detrimental effects of reward on performance: a literature review and a prediction model. In LEPPER, M. R. and GREENE, D. (Eds.), *The Hidden Costs of Reward*. Hillsdale, N.J.: Erlbaum.

OSKAMP, S. (1977). *Attitudes and Opinions*. Englewood Cliffs, N.J.: Prentice-Hall.

REISS, S., and SUSHINSKY, L. W. (1975). Overjustification, competing responses, and the acquisition of intrinsic interest. *J. Person. soc. Psychol.*, 31, 1116-1125.

SJÖBERG, L. (1982). Reward, intrinsic motivation, and performance. Paper read at a workshop in Organisational Behaviour, the Tenth Anniversary of the European Institute for Advanced Studies in Management, Lauzelle, Belgium, May 24-26.

SJÖBERG, L. (1983). Interest, achievement and vocational choice. *Eur. J. sci. Educ.*, 5, 299-307.

SJÖBERG, L., and BIEL, A. (1983). Mood, belief-value distortion and action. *Acta Psychol.*, 53, 253-270.

SVENSON, O. (1981). Are we all less risky and more skilful than our fellow drivers? *Acta Psychol.*, 47, 143-148.

TOMKINS, S. D. (1962). *Affect, Imagery, Consciousness*. New York: Springer.

TYLER, L. E. (1965). *Psychology of Human Differences*. Englewood Cliffs, N.J.: Prentice-Hall.

WEINER, B. (1979). A theory of motivation for some classroom experiences. *J. educ. Psychol.*, 71, 3-25.

(*Manuscript received* 30*th November*, 1982)

Br. J. educ. Psychol., 54, 206-213, 1984

EXPERIMENTAL MANIPULATIONS OF UNIVERSITY STUDENT MOTIVATION AND THEIR EFFECTS ON EXAMINATION PERFORMANCE

BY H. W. MARSH

(Department of Education, University of Sydney)

SUMMARY. Students (N = 416) viewed a videotaped lecture and then completed an objective examination based upon the lecture. The lecture was experimentally manipulated to vary in content coverage (the number of test questions covered) and the expressiveness with which it was delivered. Different groups of students received no external incentive to do well, were told before the lecture that they would receive money for each correctly answered question (incentive-to-learn-and-perform), or were told of the added incentive after the lecture but before the examination (incentive-to-perform). Better student performance was associated with added student incentive, better content coverage, and more lecturer expressiveness. However, the level of incentive interacted with both content coverage and expressiveness. Lecturer expressiveness had a large effect ($eta^2 = 9{\cdot}4$ per cent) when there was no added incentive, but no significant effect when an external incentive was added. Content coverage had a small effect ($eta^2 = 5{\cdot}2$ per cent) with no added incentive, a larger effect ($eta^2 = 13{\cdot}0$ per cent) for the incentive-to-perform condition, and the largest effect ($eta^2 = 26{\cdot}5$ per cent) for the incentive-to-learn-and-perform conditions. These findings suggest that lecturer expressiveness has a substantial impact when extrinsic motivation is low, and that added incentives have separate effects on motivation to learn and motivation to perform.

INTRODUCTION

THE ' Dr. Fox ' effect is defined as the overriding influence of lecturer expressiveness on students' evaluations of college/university teaching. The results of the Dr. Fox studies have been interpreted to mean that an enthusiastic lecturer can ' entice ' favourable evaluations, even though the lecture may be devoid of meaningful content, and have been cited as evidence of the invalidity of student ratings (see Marsh, in press). The general Dr. Fox paradigm (see Ware and Williams, 1975; Williams and Ware, 1976; Marsh and Ware, 1982) employs a factorial design in which the expressiveness of the lecturer and the level of content coverage (i.e., the number of substantial teaching points) are systematically varied. In the original studies conducted by Ware and Williams, a series of six lectures, all presented by the same professional actor, were videotaped. Each lecture represented one of three levels of content coverage and one of two levels of lecturer expressiveness (i.e., the expressiveness with which the actor delivered the lectures). After viewing one of the videotaped lectures, students evaluated the teaching effectiveness with a rating instrument and then completed an objective examination based upon all the teaching points presented in the high content lecture. Other researchers have also employed the Dr. Fox paradigm, sometimes considering other experimental manipulations in addition to those of expressiveness and content coverage, but the emphasis has always been on the influence which these variables have on student ratings.

Reviews and meta-analysis of the Dr. Fox research (Abrami *et al.*, 1982; Ware and Williams, 1979, 1980) have consistently concluded that differences in expressiveness explained more variance in student ratings than did differences in content coverage. However, Marsh and Ware (1982) reanalysed data from the Ware and Williams studies, and drew different conclusions. A factor analysis of their rating instrument identified five evaluation factors which varied in the way they were affected by experimental manipulations. Particularly in the condition most like the university classroom where students were given incentives to do well on the achievement test, the Dr. Fox effect was not supported in that: (1) the instructor expressiveness manipulation only affected one of five rating dimensions—ratings of instructor enthusiasm, the dimension most logically related to the manipulation, and (2) content coverage affected ratings of instructor knowledge—the factor most logically related to that manipulation. This reanalysis demonstrates the importance of considering the multidimensionality of student ratings and not assuming that any variable which affects ratings is necessarily a source of invalidity (also see Marsh, 1982, for further discussion).

Surprisingly, less emphasis in the Dr. Fox research has been placed on the impact of experimental manipulations on student achievement scores, and this is the focus of the present investigation. Abrami *et al.* (1982), in their meta-analysis of this research, reported that the effect of content coverage on achievement scores has been consistently large, while the effect of instructor expressiveness has been smaller and less consistent. Since the content manipulation consists of the number of test questions which are covered on the achievement measure, this effect is neither particularly surprising nor very interesting. However, of greater importance, is the finding that the level of lecturer expressiveness affects subsequent student performance on objective achievement examinations.

Motivation is the arousal, direction and continuance of behaviour (Anderson *et al.*, 1975). Motivation theories typically distinguish between intrinsic and extrinsic motivation. Intrinsic motivation refers to some state of the learner, and is usually considered to be an individual difference variable (e.g., achievement motivation) or is manipulated by varying the level of deprivation or arousal. Extrinsic motivation typically refers to the incentive value of the reward or reinforcement. These types of motivation and perhaps others, combine to form some level of overall motivation. For example, in the Hullian model (see Hilgard and Bower, 1966), drive level (intrinsic motivation) and incentive level (extrinsic motivation) are hypothesised to combine multiplicatively with habit strength. Ball (1977) argues that motivation is clearly multidimensional, but that most research has been done with single motivational constructs to the exclusion of others. Little research has been done in educational settings to determine whether different types of motivation can substitute for each other in obtaining an optimal motivation level. In the classroom, teachers motivate students by stimulating interest (intrinsic motivation) and by providing rewards for the desired behaviours (extrinsic motivation). In the present investigation these two forms of motivation are operationalised as the level of expressiveness and enthusiasm by the instructor and as the level of monetary reward students are offered for correctly answering examination questions. The purpose of the study is to determine how these factors, along with the amount of content that is presented to students, combine to determine performance on an objective examination. This study is based upon a further analysis of the investigation presented by Marsh and Ware (1982), where the influence of these experimental manipulations on students' evaluations of teaching effectiveness was emphasised.

METHOD

Participants and design

This study is based upon further analysis of two previously published studies by Ware and Williams (1975) and Williams and Ware (1976). The data are described in more detail by Marsh and Ware (1982). Students (N = 416) viewed one of six videotaped lectures, completed an 18-item survey designed to evaluate teaching effectiveness, and took a 26-item multiple choice examination. The six videotaped lectures were all on the biochemistry of learning and memory and presented by the same professional actor, but varied systematically in terms of content coverage and lecturer expressiveness. Verbatim scripts, based upon different numbers of substantive teaching points, were used to control the content coverage (high content = 26 points; medium content = 14 points; low content = 4 points). Filler material was used to approximately equate the length of each lecture, which was about 21 minutes. Lecturer expressiveness was varied by instructing the professional actor who delivered the lectures to give each in a high or low expressive manner. Levels of expressiveness were associated with differences in enthusiasm, friendliness, charisma, humour, and ' personality '. Findings based upon the students' evaluations of teaching effectiveness are not considered in this study, but are presented elsewhere. Those results indicate that the instructor expressiveness manipulation had a huge effect on student ratings of ' Instructor Enthusiasm ', accounting for more than 25 per cent of the variance in that rating factor, and substantially smaller effects in student ratings of ' Clarity/Organisation ', ' Instructor Concern ', ' Instructor Knowledge ', and ' Learning Stimulation ' (the manipulation accounted for no more than 8 per cent of the variance in any of these other factors). Thus, the experimental manipulation of instructor expressiveness was quite effective.

The 26-item, multiple-choice examination consisted of 11 knowledge-level and 15 comprehension-level items. Each of the 26 teaching points in the high content lecture was tested by a single item, and students in the medium and low content conditions were tested on materials not covered in the videotaped lecture which they viewed.

Students (N = 205) in the first study (Ware and Williams, 1975) were undergraduate and graduate students who were enrolled in one of three sections of an undergraduate general studies course at a midwestern university in the United States. All students who attended the class on the day of the experiment participated in the study. These students in the first study were given little or no incentive to learn the materials or to perform well on the examination, and the outcome of their performance had no effect on their course grade. Students in the second study (Williams and Ware, 1976) were enrolled in one of 12 sections of an undergraduate course in educational psychology at a midwestern university in the United States. Again, all students who attended the class on the day of the experiment participated in the study, and their performance in the experiment had no effect on their course grade. Students in the second study were given one of two levels of monetary incentive that was based upon test performance. At both levels of incentive, students were paid $1 for participating in the study, and an additional 5¢ for each item on the achievement test that they answered correctly. In one condition (incentive-to-learn-and-perform), students (N = 104) were told of the added incentive *before* viewing the lecture or taking the examination. In the other condition (incentive-to-perform), students (N = 107) were told of the added incentive *after* viewing the lecture but before taking the examination. Thus, across the two studies, there are three levels of incentive: no external incentive (1975 study), incentive to perform better after the opportunity to learn had occurred, and incentive to learn and perform better.

In the first study, students from a special studies course were randomly divided into six experimental groups. In the second study, 12 intact sections of an educational psychology course constituted the experimental groups. *Within* each of the studies there was random assignment and the experimental groups were similar in terms of age, sex, GPA, and a prior interest in the lecture topic. Students in the 1975 study, compared with those in the 1976 study: (*a*) were from a general studies course rather than an educational psychology course; (*b*) were more likely to be freshmen and were, on the average, a year younger; and (*c*) indicated less initial knowledge and interest in the content of the lecture. However, student self-reports of academic major, age, year in school, prior knowledge, and prior interest were not empirically correlated with either examination performance or student ratings of teacher effectiveness.

Analysis

By combining the six experimental groups from the 1975 study with the 12 from the 1976 study, a total of 18 groups was obtained. A series of group comparisons using analysis of variance (ANOVA) was conducted with the commercially available BMDP statistical package (Dixon and Brown, 1979, pp. 540-580). In the first analysis, the 18 groups were compared with a 3 (level of content) × 2 (level of expressiveness) × 3 (level of incentive) ANOVA. Then, a two-way ANOVA (content × expressiveness) was conducted separately for each of the three incentive conditions. Finally specific group means were compared by using Newman-Keuls tests to assist in the interpretation of significant F-ratios. Eta squared was used to estimate the magnitude of the different effects.

RESULTS AND DISCUSSION

The three-way ANOVA (see Table 1) indicates that the effects of incentive level (eta^2 = 13·0 per cent) and content coverage (eta^2 = 10·7 per cent) are largest, but that lecturer expressiveness is also statistically significant (eta^2 = 1·7 per cent, $P < 0{\cdot}01$).

Newman-Keuls tests demonstrated that differences between each of the three incentive groups, and between each of the three content groups are statistically significant (all $P < 0{\cdot}01$). Inspection of the group means (see Figure 1) shows that better examination performance is associated with higher levels of expressiveness, better content coverage, and added student incentive. Furthermore, telling students of the added incentive before the lecture (incentive-to-learn-and-perform condition) produces better examination performance than telling them before the examination (incentive-to-perform conditions).

Most of the variance due to the experimental manipulations can be explained in terms of the main effects, but the incentive manipulation does interact with expressiveness (eta^2 = 1·2 per cent, $P < 0{\cdot}5$) and content coverage (eta^2 = 1·8 per cent, $P < 0{\cdot}05$). In order to explore these interactions, a series of two-way ANOVAs was conducted

FIGURE 1

EFFECTS OF INCENTIVE ON EXAMINATION PERFORMANCE

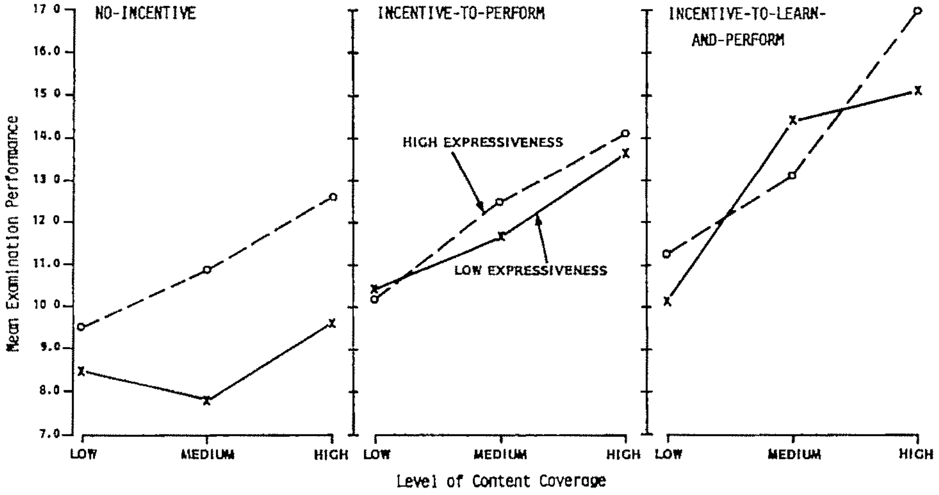


TABLE 1

THE INFLUENCE OF DIFFERENT EXPERIMENTAL MANIPULATIONS ON STUDENT EXAMINATION PERFORMANCE

	Across all three incentive groups				No-incentive group				Incentive-to-perform group				Incentive-to-perform-and-learn group			
Source	DF	MS	F	Eta^2	DF	MS	F	Eta^2	DF	MS	F	Eta^2	DF	MS	F	Eta^2
Incentive (A)	2	473·15	37·7**	13·0%												
Expressiveness (B)	1	119·49	9·5**	1·7%	1	250·76	22·2**	9·4%	1	5·19	0·35		1	9·07	0·7	
Content (C)	2	381·84	30·4**	10·7%	2	69·09	6·1**	5·2%	2	114·05	7·63**	13·0%	2	234·44	18·5**	26·5%
A × B	2	42·72	3·4*	1·2%												
A × C	4	32·60	2·6*	1·8%												
B × C	2	12·31	1·0		2	16·90	1·5		2	3·09	0·21		2	24·23	1·91	
A × B × C	4	15·95	1·3													
Error	398	12·55			199	11·28			101	14·95			98	12·67		

* $P < 0·05$ ** $P < 0·01$

Note: A 3 (Incentive) × 2 (Expressiveness) × 3 (Content) ANOVA was performed across all 18 experimental groups, and then three two-way ANOVAs were performed separately for each of the three incentive conditions. Eta^2, the effect size, is defined as (SSeffect/SStotal) × 100%.

separately at each level of incentive (see Table 1 and Figure 1). Lecturer expressiveness ($eta^2 = 9{\cdot}4$ per cent) has a large effect when there is no added incentive, but fails to reach statistical significance in either of the other two incentive conditions. Content coverage has a small effect ($eta^2 = 5{\cdot}2$ per cent) when there is no added incentive, a larger effect in the incentive-to-perform condition ($eta^2 = 13{\cdot}0$ per cent), and its largest effect in the incentive-to-learn-and-perform conditions ($eta^2 = 26{\cdot}5$ per cent). Newman-Keuls tests were used to compare differences due to incentive levels separately for each of the three content conditions. At the low level of content, the difference between the incentive-to-perform and the incentive-to-learn-and-perform conditions failed to reach statistical significance, but all other comparisons did.

The effect of the different levels of incentive shows that by offering students an added incentive their examination performance will improve. Furthermore, the significant difference between the incentive-to-perform and incentive-to-learn-and-perform conditions suggests that the added incentive has separate effects on the motivation to learn and the motivation to do well on the test. Students who were told of the added incentive just before taking the examination presumably had no opportunity to improve their learning of the materials, so their improvement over the no-incentive conditions appears to be a performance effect rather than better learning. Students who were told of the added incentive before viewing the lecture had the same incentive to perform as students in the incentive-to-perform conditions, thus their improved performance appears to be a function of better learning as well as their motivation to perform well. The fact that students also completed the teacher rating form between viewing the lecture and taking the examination makes it more plausible that ' unlearned ' information from the lecture was unavailable to students when they took the examination. However, the difference between the incentive-to-perform and incentive-to-learn-and-perform conditions is likely to have been larger if there had been a longer delay between the lecture and the examination.

The main effect of content coverage does not have important implications. Indeed, if varying the number of examination questions that are covered in the lecture does not substantially affect examination performance, then the validity of the experimental manipulation is questionable. However, the interaction between content coverage and level of incentive is more interesting. When students have no external incentive, covering additional content (i.e., 4 of 26 test items *vs.* all 26 items) has only a small effect on examination performance (9 *vs.* 11 questions correctly answered). Students appear unable or unwilling to process the additional information in the high content lecture. In contrast, for the incentive-to-learn-and-perform condition, the same content manipulation has a much larger effect on examination performance (10·7 *vs.* 16 questions correctly answered). Thus, with added incentive to learn and perform better, students are more willing to process the additional information in the high content lecture. In the incentive-to-perform condition, the content manipulation has a moderate effect (10·3 *vs.* 14 questions correctly answered). It is only when the information demands are higher that knowing about the added incentive before the lecture produces significantly better performance. These results suggest that students are willing to process more information when their incentive is higher, that level of incentive has a small effect when information demands are low, but that incentive has a much larger effect when information demands are higher.

The effect of lecture expressiveness is particularly important. Educators often argue for the importance of instructor characteristics such as expressiveness and enthusiasm, but demonstrations are typically limited to correlational designs that employ some form of teacher ratings. Within this framework causal relationships cannot be shown, and these characteristics are often classified as relatively unchangeable personality characteristics which have a dubious role in the definition of effective teaching. Indeed, the original purpose of the present set of studies was to explore the contention that lecturer expressiveness biases student ratings of teacher effectiveness in such a way that teachers can ' seduce ' students into giving them favourable evaluations by merely presenting materials in a highly expressive and enthusiastic manner (which is not the case—see Marsh and Ware, 1982). In contrast to most other studies, the present results show that more expressively presented lectures *cause* better examination performance in a study where there was random assignment to treatment conditions and lecturer expressiveness was experimentally manipulated.

The expressiveness-by-incentive interaction is also quite important. When students have no external incentive, the effect of lecturer expressiveness on examination performance is substantial—even more important than the number of examination questions that are covered in the lecture. If extrinsic motivation is low, students may need to have lectures presented in an enthusiastic and expressive manner. In contrast, when an external incentive is added, expressiveness has little effect and content coverage has more to do with how well students do on the examination. If extrinsic motivation is high, then lecturer expressiveness may be less important and content coverage more important.

Findings from this study should not be translated too literally into the classroom setting. The " learning environment " was limited to a single 20-minute lecture. The examination was administered almost immediately after the lecture was completed, and stressed memory rather than higher cognitive processes. The design of the study afforded students no time for homework, outside reading, or additional study. The motivation effects of added incentives and lecturer expressiveness are likely to be much more complicated in actual classrooms. In particular, 5 cents per right answer seems a remarkably small amount to produce such a large shift in achievement scores after a lecture presented in an unexpressive manner. (We suspect that it was not the monetary value *per se* which was the critical feature, but rather the knowledge that the tests would be scored and evaluated, and the results would be returned. Thus, " points " may have served as well.) Nevertheless, the findings do have intuitive appeal, are consistent with results from human learning studies (e.g., the distinction between learning and performance), and seem to have a broad application. Furthermore, the general procedure whereby specific characteristics of the lecture are systematically manipulated can be applied to a wide range of variables while still providing a level of control which is unusual in studies of teaching effectiveness.

REFERENCES

ABRAMI, P. C., LEVENTHAL, L., and PERRY, R. P. (1982). Educational seduction, *Rev. educ. Res.*, 52, 446-464.

ANDERSON, S. B., BALL, S., MURPHY, R. T., and ASSOCIATES (1975). *Encyclopedia of Educational Evaluation.* San Francisco: Jossey-Bass.

BALL, S. (1977). *Motivation in Education.* New York: Academic Press.

DIXON, W. J., and BROWN, M. B. (1979). *BMDP Biomedial Computer Programs.* Los Angeles: University of California Press.

HILGARD, E. R., and BOWER, G. H. (1966). *Theories of Learning.* New York: Appleton-Century-Crofts.

MARSH, H. W. (1982). SEEQ: A reliable, valid, and useful instrument for collecting students' evaluations of university teaching. *Br. J. educ. Psychol.*, 52, 77-95.

MARSH, H. W. (in press). Students as evaluators of teaching. In HUSEN, T., and POSTLETHWAITE, T. N. (Eds), *International Encyclopedia of Education: Research and Studies.* Oxford: Pergamon Press.

MARSH, H. W., and WARE, J. E. (1982). Effects of expressiveness, content coverage, and incentive on multidimensional student rating scales: new interpretations of the Dr. Fox effect. *J. educ. Psychol.*, 74, 126-134.

WARE, J. E., and WILLIAMS, R. G. (1975). The Dr Fox effect. A study of lecture expressiveness and ratings of instruction. *J. medic. Educ.*, 5, 149-156.

WARE, J. E., and WILLIAMS, R. G. (1979). Seeing through the Dr. Fox effect: a response to Frey. *Instruct. Eval.*, 3, 6-10.

WARE, J. E., and WILLIAMS, R. G. (1980). A reanalysis of the Doctor Fox experiments. *Instruct. Eval.*, 4, 15-18.

WILLIAMS, R. G., and WARE, J. E. (1976). Validity of student ratings of instruction under different incentive conditions: a further study of the Dr. Fox effect. *J. educ. Psychol.*, 68, 48-56.

(*Manuscript received 22nd May*, 1982)

Br. J. educ. Psychol., 54, 214-220, 1984

TRAINING TEACHERS IN COUNSELLING SKILLS: THE EFFECTS OF MICROCOUNSELLING

BY O. D. W. HARGIE
(*Social Skills Centre, Ulster Polytechnic, Northern Ireland*)

SUMMARY. This paper examines the effects of a programme of microcounselling upon the training of in-service teachers in specific counselling skills. It is argued that differences exist between the patterns of teacher questions common in classrooms and those appropriate in counselling, and that similar differences can be found between the proportion of teacher- and pupil-talk in both settings. It was found that, following microcounselling, teachers significantly reduced the number of closed questions used, and also effected a significant increase in the total amount of pupil talk. In addition, teachers increased their use of open questions, and decreased their overall talk-times, although neither of these results reached the 0·05 level of significance. It is concluded that microcounselling is a promising, if still evolving, technique for training teachers in the skills of counselling.

INTRODUCTION

MICROCOUNSELLING, as a technique for training counsellors in the skills of counselling, has been in operation for some 15 years (Ivey *et al.*, 1968), and during this time, a considerable body of research has been accumulated into determining the relative efficacy of this technique. The findings from this research have been generally positive, and it has been shown that, following microcounselling, trainees demonstrate improvements on a range of relevant communication skills (Ivey and Authier, 1978). Furthermore, when compared to alternative methods of counsellor training, " microcounselling has also been discovered to have greater potency than the experiential component of traditional training approaches " (Dickson, 1981, p. 72). Given this state of affairs, it is not surprising that microcounselling has been employed in the training of a wide range of professionals including social workers, doctors, nurses, policemen, psychologists, clergymen and careers advisers (Hargie and Saunders, 1983).

Whilst microcounselling, at a general level, has been shown to be effective, much more specific research is needed in order to ascertain exactly to what extent, with which specific skills, and in relation to which groups of trainees, it is effective. In the present study the focus is upon a group of in-service teachers enrolled on a guidance and counselling course at Ulster Polytechnic. In this case, the teachers needed to develop and refine their repertoire of skills appropriate to offering a guidance and counselling service to the pupils in their schools. In order to do so, however, it is often necessary for them to unlearn, or at least re-focus, some of their existing teaching skills. Both teaching and counselling involve a range of interpersonal skills necessary for effective practice. While some of these skills may appear to be very similar, the focus is often quite different. Thus, for example, teachers tend to utilise overt verbal reinforcement in order to encourage pupil participation and learning; counsellors, on the other hand, tend to employ more non-verbal reinforcers and less direct verbal reinforcement (Hargie *et al.*, 1981). Similar differences in focus can be found in skills such as opening and closing, explanation, non-verbal communication, listening and questioning.

The main reason for these differences lies in the nature of the two activities. Teaching involves interacting with, and imparting knowledge to, a group of individuals in a situation in which the teacher often is the main focus of attention, and initiator of activities. Counselling, on the other hand, is primarily a dyadic encounter in which the counsellee, or client, is expected to initiate and be the central focus. The changes in behaviour pattern necessary to adapt from being a teacher in control of classroom activities, to a counsellor willing to play a more non-directive role may well cause problems for some teachers. This study investigates the effects of such a transition on two specific interpersonal areas, namely the use of questions, and teacher/pupil talk ratios.

Research in the classroom has illustrated that over 60 per cent of teacher questions are of a closed nature designed to test pupils' recall of factual material, with only 21 per cent of questions asked leaving pupils open to answer in a thoughtful, evaluative fashion (Turney *et al.*, 1976). As Hargie *et al.* (1981) point out, these proportions are " indicative of the examination oriented curriculum in schools, which encourages teachers to emphasise memory skills " (p. 71). Nevertheless, a number of educationists have urged teachers to make more use of open questions to encourage critical thinking in pupils (Merwin and Schneider, 1973; Dillon, 1982). Research findings support such a viewpoint. Rousseau and Redfield (1980) in reviewing a total of 20 studies in this field conclude that " gains in achievement over a control group may be expected for groups of children who participate in programmes where teachers are trained in questioning skills . . . gains are greatest when more higher cognitive questions are used during instruction " (p. 52). It is not surprising, therefore, that in dyadic counselling contexts, the advantages of open, as opposed to closed, questions have been emphasised by, amongst others, Ivey (1971), Kaplan (1976), and King (1972).

The relative advantages of open questions include allowing the respondent to express opinions, attitudes, thoughts and feelings more fully. They also place fewer restrictions upon the respondent, who thereby has a higher degree of control over the interaction, and can determine to a greater extent what is to be discussed.

Research comparing the use of open and closed questions in counselling has found that open questions are more effective in promoting interviewee self-disclosures (Beharry, 1976), in producing more accurate responses (Dent and Stephenson, 1979), and increasing perceived counsellor empathy (Rennie *et al.*, 1978). Furthermore, Dohrenwend (1965) found that responses to open questions were approximately three times longer than responses to closed questions. This latter finding is of importance to teachers, since it has been found that teachers will often talk for as much as two-thirds of the total talk time in classrooms, leaving pupils to contribute the remaining one-third (Hargie, 1980). When this pupil talk proportion is divided between up to 30 pupils per class, then it is obvious that each particular pupil does not contribute significantly to the total classroom talk time. It is probable that the preponderance of closed questions asked by teachers is one of the causes of the overall reduced level of pupil, as opposed to teacher, talk.

The significance of these classroom research findings for the training of teacher counsellors is fairly obvious. In the guidance and counselling situation, the teacher counsellor should encourage the pupil to talk freely, and at length, about any problem or difficulty he or she may have. As a result, the counsellor should make use of more open, as opposed to closed, questions, and should generally encourage the client to do most of the talking (Crites, 1974). These behavioural objectives are, however, the converse of the teacher's classroom interaction style and, therefore, require the teacher to make major behavioural changes from the classroom to the counselling context.

In order to facilitate such changes, some form of training is usually required, and this training has increasingly taken the form of microcounselling. In this study a group of in-service teachers, enrolled on a one-year, full-time Diploma in Guidance and Counselling at Ulster Polytechnic, were given a programme of microcounselling, and changes in questioning behaviour, and in teacher/pupil talk times were monitored pre- and post-training. It was predicted that, following microcounselling, the teachers would significantly increase their use of open questions and decrease their use of closed questions, and that a significant increase in pupil talk-time and decrease in teacher talk-time would occur.

METHOD

Subjects

Seven teachers (five male, two female), all with several years classroom experience and currently teaching in secondary education (11-18 years). Their teaching expertise ranged from arts subjects to science and physical education, and they were all seconded on full pay to attend a one-year full-time course of training in guidance and counselling. Prior to attending the course their counselling expertise was limited to informal, *ad hoc* consultation with pupils, and the training course was designed to equip them with skills to enable them to implement a guidance and counselling service in school.

Training

The Diploma course is organised in two parts, each lasting one semester (14 weeks). Students undertake modules on human and vocational development, microcounselling, understanding research, methods of assessment, principles and practice of guidance and counselling, sociology, education and current issues. The microcounselling programme, therefore, amounted to some 12 per cent of the total year's work. This programme was taught every Monday and Thursday for the first semester from 9.15 to 12.15. The first session of the week was given over to theory, and the second session was a practical (see Figure 1).

A typical theoretical session would include a lecture supplemented by handouts in which the rationale for the skill in question is offered, an outline of related research is given and students are presented with a description of the behavioural elements of each skill. This is followed by initial analyses of videotaped exemplars of the skill in guidance and counselling encounters, and a general discussion on the relative advantages and disadvantages of utilising the skill in guidance and counselling.

FIGURE 1

THE MICROCOUNSELLING PROGRAMME

Week	Lecture	Practical
(1)	The Social Skills Model	Pre-tape (pupils)
(2)	Non-verbal communication	1-1 interviews (pupils)
(3)	Questioning	1-1 interviews (pupils)
(4)	Reinforcement	1-1 interviews (pupils)
(5)	Reflecting and Paraphrasing	1-1 interviews (peers)
(6)	Listening and Sustaining	1-1 interviews (peers)
(7)	Group Interaction	Peer group exercise
(8)	Integration of Skills	1-1 interviews (pupils)
(9)	Set Induction/Closure	Group work (pupils)
(10)	Group Discussion	1-1 interviews (pupils)
(11)	Explaining	1-1 interviews (pupils)
(12)	Plenary and integration	Post-tape (pupils)

The practical session usually involved each student practising the skill under consideration in a ' scaled-down ' interview lasting up to 15 minutes. The interviewees were all pupils from neighbouring secondary schools who came to the Polytechnic to be interviewed. Throughout the programme a range of pupils of differing ages (15-18 years) and sexes participated in microcounselling. For practice of the skills of reflecting and listening trainees role-played the part of pupils with emotional problems. This served two main purposes: firstly, it encouraged teachers to put themselves in the role of pupils and gain further insight into how they might feel in certain situations, and, secondly, it facilitated the analysis of difficult, personal, problems which, for ethical reasons, ' real ' pupils were not expected to present during microcounselling, especially since all interviews were video-recorded. These video-recordings took place in the Social Skills Training Centre (see Hargie, 1980, for a full description) by means of a hidden remotely-controlled camera. Thus, although both interviewer and interviewee knew they were being recorded, no CCTV equipment was observable in the interview room and no one else was present during interviews. A tutor monitored the interview on a TV screen in an adjoining room, which served as a playback room during feedback sessions. The group of seven was sub-divided into two groups of three and four, and each sub-group viewed and discussed with a tutor all of the videotapes of each student in that particular group.

Procedure

The pre-tape was made at the beginning of the course after an introductory lecture when trainees were told of the nature of social skills training. There were no specific skills taught at this stage and trainees were instructed to interview a school-leaver from a nearby school on the general area of family background, school and career choice. As a general guide, trainees were shown a typical guidance interview. The post-tape was made upon completion of microcounselling, when trainees had received the full range of training and were attempting to integrate what they had learned over the previous three months. On both occasions trainees interviewed 16-year-old boys from the same form of a local secondary school, with each trainee interviewing a different pupil. Thus, a total of 14 pupils was involved in pre- and post-tapes—one pupil for each trainee on both occasions. The fact that all pupils were from the same form was intended to provide a degree of control for IQ. No control was made for personality.

Each videotape was made by recording the trainee interviewing a pupil. The pupil entered the interview room where he was received by the interviewer and offered pre-arranged seating, comprising two soft chairs positioned in such a way that both participants could be clearly observed on CCTV. The signal to begin, and terminate, the interview was given by a warning light situated behind the interviewee. Each pre- and post-tape should have lasted a total of 15 minutes. However, in the pre-tape two trainees terminated their interviewees after 13 minutes, and, in order to ensure uniformity, the analysis of tapes therefore comprised the first 13 minutes of each pre- and post-tape.

All of the tapes were analysed on two occasions. The dimension of analysis for the first section was the distinction between open and closed questions and the total frequency of occurrence of each type was noted. For the second section two cumulative stop watches were used to count total pupil and trainee talk-times.

RESULTS

A comparison of pre- and post-test scores indicates a significant decrease in the mean score for closed questions from 49·1 to 37, ($P < 0.05$), and an increase, though non-significant, in the mean score for open questions from 5·7 to 6·3. Thus, both of these results are in the predicted directions (see Table 1). It should be noted, however, that two of the seven trainees (' A ' and ' E ') actually demonstrated a decrease in open questions, and another trainee (' F ') used exactly the same number, from pre- to post-test. Thus, the microcounselling programme seems to have had no consistent effect on this variable. One trainee (' B ') increased the number of closed questions pre- to post-test while the remaining six decreased their use of closed questions.

Changes in teacher-, and pupil-, talk times were in the predicted directions (see Table 2). Total pupil talk increased significantly from a mean of 258·6 secs in the pre-test to 382·9 secs in the post-test ($P < 0{\cdot}05$). Correspondingly, a non-significant decrease in teacher talk occurred from pre- to post-tests with mean scores of 460 secs and 342 secs respectively. It is interesting to note that one trainee (' D ') showed an increase in talk time from pre- to post-test, and the related pupil talk time decreased accordingly. In the post-test, the mean scores indicate that, overall, pupils talked slightly more than trainees, whereas in the pre-test the trainees talked markedly more than pupils.

TABLE 1

CHANGES IN OPEN AND CLOSED QUESTIONS FOLLOWING MICROCOUNSELLING

	Open questions		Closed questions	
Teacher	Pre	Post	Pre	Post
A	10	6	50	27
B	5	8	55	57
C	2	5	49	45
D	1	4	42	41
E	10	7	48	28
F	6	6	46	34
G	6	8	54	27
x	5·7	6·3	49·1	37*
SD	3·5	1·5	4·5	11·38

* Correlated *t*-test: $P < 0{\cdot}05$

TABLE 2

CHANGES IN TEACHER AND PUPIL TALK-TIMES FOLLOWING MICROCOUNSELLING (IN SECONDS)

	Teacher		Pupil	
	Pre	Post	Pre	Post
A	438	255	289	484
B	482	363	258	353
C	619	332	101	388
D	321	419	342	302
E	524	266	207	454
F	544	518	205	220
G	292	242	408	479
x	460·0	342·1	258·6	382·9*
SD	119·9	101·9	101·2	99·7

* Correlated *t*-test: $P < 0{\cdot}05$

DISCUSSION

This study indicates that microcounselling is a useful training method with which to equip teachers with skills appropriate to the dyadic helping process. At the same time, it would seem that certain behaviours may be more amenable to this form of training than others. Previous research has indicated that teachers need to develop their use of open, higher-order questions, in classroom settings, and has recommended the introduction of training techniques such as microteaching to meet this need (Hargie, 1978). The finding, in this study, that teachers did not significantly increase their use of open questions suggests that more attention needs to be paid to the development of this technique in teacher counsellors. Since teachers tend to use few open questions in their classroom settings, it is therefore likely that more time needs to be devoted to the practice of this skill in microcounselling programmes. This may involve the use of several simulated interview sessions where teachers can practise open questions, or the introduction of follow-up questioning practicals throughout a microcounselling programme.

Teachers in this study were able to significantly reduce the overall number of closed questions used, and this suggests that where mastery of a skill already exists, appropriate usage can be more readily inculcated.

The nature of the interviews in the present investigation needs to be considered in relation to the differential use of types of questions. In both the pre- and post-tape interviews, trainees were faced with a pupil whom they were meeting for the first time. Thus, a certain proportion of the interview was taken up by " fact-finding ", and, for this purpose, closed questions are both acceptable and effective (Benjamin, 1974). Although trainees were encouraged to use open questions for fact-finding, the pressure of time, coupled with existing behavioural style, resulted in a preponderance of closed questions, which in turn served to give trainees more control over the direction and content of the interview. In addition, there was a time delay of nine weeks between the analysis and practice of the questioning skill and the post-tape. This delay, coupled with the intervening emphasis on a range of other skills, may have helped to negate any learning regarding increases in open questions.

In relation to teacher and pupil talk-times, microcounselling had a marked effect on teachers, in terms of reducing the amount of time spoken by the teacher and increasing the amount of time spoken by pupils. Given the importance of allowing clients to participate, and indeed take the initiative, in counselling, it is vital that counsellors adopt techniques which facilitate this process. Yet, for teachers accustomed to taking the leading role in classrooms, this is often a difficult approach to master. Indeed, a comment frequently expressed by teachers, during microcounselling practicals with pupils, was " I kept wanting to get in there ", and, in the early stages of training, interruptions by the teacher counsellor were quite frequent. By the end of the programme, however, the teachers had managed to reduce their overall amount of talk, and to significantly increase the amount of pupil talk during interviews.

Overall, therefore, the microcounselling approach adopted with these in-service teachers proved to be fairly effective in terms of producing stylistic changes in performance, away from a teacher-centred classroom approach towards a client-centred counselling approach. Since the number of subjects employed in this study was small, however, it is difficult to make any firm generalisations from these results. Much more research is needed before any definite conclusions can be made regarding the overall effectiveness of microcounselling with in-service teachers. It may be that some teaching behaviours will be more resilient to change than others, or that approaches other than microcounselling may be employed to effect changes in certain behaviours. Furthermore, long-term studies are needed in order to investigate how, and to what extent, changes in behaviour following microcounselling are maintained over time.

ACKNOWLEDGMENTS.—I would like to thank Peter McAllister who undertook the analysis of the videotapes, and Pat Penhallurick, Rob Millar, David Coffey and Pauline Irving from the Guidance and Counselling Division, Ulster Polytechnic who helped organise and implement the microcounselling programme.

REFERENCES

BEHARRY, E. (1976). The effect of interviewing style upon self-disclosure in a dyadic interaction. *Dissert. Abstr. Internat.*, 36, 4677B.
BENJAMIN, A. (1974). *The Helping Interview.* Boston: Houghton Mifflin.
CRITES, J. (1974). Career counselling: a review of major approaches. *The Counselling Psychologist*, 4, 3-38.
DENT, H., and STEPHENSON, G. (1979). Experimental study of the effectiveness of different techniques of questioning child witnesses. *Br. J. soc. clin. Psychol.*, 18, 41-51.
DICKSON, D. (1981). Microcounselling: an evaluative study of a programme. Unpublished Ph.D. thesis, Ulster Polytechnic, N. Ireland.
DILLON, J. (1982). The multidisciplinary study of questioning. *J. educ. Psychol.*, 74, 147-165.
DOHRENWEND, B. (1965). Some effects of open and closed questions on respondents' answers. *Human Organ.*, 24, 175-184.
HARGIE, O. (1978). The importance of teacher questions in the classroom. *Educ. Res.*, 20, 99-102.
HARGIE, O. (1980). An evaluation of a microteaching programme. Unpublished Ph.D. thesis, Ulster Polytechnic, N. Ireland.
HARGIE, O., and SAUNDERS, C. (1983). Training professional skills. In DOWRICK, P., and BIGGS, S. (Eds), *Using Video.* Chichester: Wiley.
HARGIE, O., SAUNDERS, C., and DICKSON, D. (1981). *Social Skills in Interpersonal Communication.* London: Croom Helm.
IVEY, A. (1971). *Microcounselling: Innovations in Interviewing Training.* Springfield, Ill.: C. C. Thomas.
IVEY, A., and AUTHIER, J. (1978). *Microcounselling: Innovations in Interviewing, Counselling, Psychotherapy and Psychoeducation.* Springfield, Ill.: C. C. Thomas.
IVEY, A., NORMINGTON, C., MILLER, C., MORRILL, W., and HAASE, R. (1968). Microcounselling and attending behaviour: an approach to prepracticum counsellor training. *J. counselling Psychol.*, 15, 1-12.
KAPLAN, E. (1976). Effective interviewing. *Special Libraries*, 67, 63-67.
KING, G. (1972). Open and closed questions: the reference interview. *RQ-Reference and Adult Services Division*, 12, 157-160.
MERWIN, W., and SCHNEIDER, D. (1973). The use of self-instructional modules in the training of social studies teachers to employ higher cognitive level questioning strategies. *J. educ. Res.*, 67, 13-18.
RENNIE, D., BURKE, H., and TOUKMANIAN, S. (1978). Counsellor communication style as a determinant of rater-perceived empathy. *Canadian Counsellor*, 12, 235-242.
ROUSSEAU, E., and REDFIELD, D. (1980). Teacher questioning. *Evaluation in Education*, 4, 51-52.
TURNEY, C., OWENS, L., HATTON, N., WILLIAMS, G., and CAIRNS, L. (1976). *Sydney Micro Skills: Series 2 Handbook.* Sydney: Sydney University.

(*Manuscript received* 14*th September*, 1983)

RESEARCH NOTES

Br. J. educ. Psychol., 54, 221-222, 1984

KINDERGARTEN TASK MANAGEMENT AND GRADE 2 PERCEPTIONS OF TEACHER-LEARNER ROLES

By B. CHRIS NASH

(*Ontario Institute for Studies in Education*)

Summary. A double-blind follow-up study of Grade 2 children showed that those who had experienced a kindergarten programme stressing personal task-management skills were more likely to perceive themselves as responsible for their own learning than those from teacher-directed kindergartens.

INTRODUCTION

An earlier article (Nash, 1979) described strategies for teaching 4- and 5-year-old children how to plan classroom learning activities. Immediate effects included the children's perception of being in control of their own learning. Recently a study was done of some children from such programmes two years later.

Two Grade 2 classes were located in which about half of the children had had the task management kindergarten programme (known as Learning Environment (LE)) and the rest had not. Neither Grade 2 teachers nor the investigator knew which of the children were LE graduates. After kindergarten, no special attempt had been made to encourage pupils to participate in decisions about their learning. The LE approach stresses the choice and pre-planning of learning activities by the child, with the teacher monitoring task completion and discussing goal attainment with the pupils.

PRESENT STUDY

To assess whether pupils understood or believed in their own responsibilities for successes and failures, 41 children were interviewed. Of these, 22 were LE graduates.

The investigator approached the pupils as they worked at arithmetic exercises and asked them about the task. She commented on how well the child had done on one page and asked to look at more of the work. She would then find a page where there were one or two (or more) mistakes and ask the child " What happened here? You got everything right here and here, but on this page you got this (these) wrong? Can you remember why? "

Most of the children responded by offering explanations which included assignment of blame for less than acceptable performance. In all but five cases this blame was placed unambiguously on either the teacher or on the pupil. " [The teacher] didn't explain it right." " She forgot I was away last week." " She didn't read [my work]." or " I missed the lesson. Then I forgot to tell Miss ——." " I didn't understand the first time. But I got it the second time. See? " " I make my threes look like twos sometimes."

When the 37 classifiable responses were checked against the children's kindergarten training in task planning, it was found that the acceptance of personal responsibility for inadequate work was associated with earlier training in task management. Details are presented in Table 1.

TABLE 1

Grade 2 Pupils' Perception of Responsibility for Errors and Kindergarten Training in Task Planning

Kindergarten experience	Personal responsibility	Teacher responsibility	Total
Task planning	16	4	20
No task planning	5	12	17
Total	21	16	37

$x^2 = 8{\cdot}67$ df = 1 P = 0·005

Further work on this topic is anticipated.

REFERENCE

NASH, B. C. (1979). Kindergarten programmes and the young child's task orientation and understanding about time scheduling. *Br. J. educ. Psychol.*, 49, 27-38.

(*Manuscript received* 3rd *August*, 1983)

Br. J. educ. Psychol., 54, 223-227, 1984

THE INFLUENCE OF CHRISTIAN NAME STEREOTYPES ON THE MARKING OF CHILDREN'S ESSAYS

BY P. G. ERWIN
(*Manchester Polytechnic*)
AND A. CALEV
(*Manchester University*)

SUMMARY. It is argued that previous name stereotype research has been one-sided in focusing solely on stimulus characteristics. This study attempted to improve this situation by taking into account the christian name stereotype of both the subject and stimulus figure. Subjects marked six essays each supposedly written by a different stimulus figure. A 2 × 2 × 2 × 3 ANOVA (male *vs.* female subjects × attractively named *vs.* unattractively named subjects × male *vs.* female stimulus figures × attractively named *vs.* unattractively named *vs.* unnamed stimulus figures) showed only two significant effects, for attractiveness of the subject's name ($P = 0{\cdot}031$) and for attractiveness of the name of the stimulus figure ($P = 0{\cdot}005$). This latter effect showed a significant linear trend ($P < 0{\cdot}005$) with unattractively named stimulus figures receiving lower marks than unnamed stimulus figures who in turn received lower marks than attractively named stimulus figures. The results were related to the general literature on social relationships and the need for further integrative research was stressed.

INTRODUCTION

Since the 1940s there has been a great deal of research on the consequences—both positive and negative—of christian name stereotyping. People appear to develop strong emotional feelings for their own name, which may in turn affect their judgments and evaluations of others (Strunk, 1958). Garwood (1976) has shown that subjects with attractive names tend to have much lower levels of inner conflict and much higher levels of expected success than their counterparts with unattractive names. Name stereotypes may also influence the form and satisfaction an individual experiences in his personal and social relationships (Buchanan and Bruning, 1971; Lawson, 1971; McDavid and Harari, 1973). Thus, it is evident that name stereotypes exert their influence in two major ways. First, they affect one's self image and, via this factor, influence social relations (for a review of the role of self esteem in social and personal relationships see Stroebe, 1977). Secondly, people's reactions to an individual and his or her accomplishments will be directly influenced by his or her name. Both these factors may contribute to the ubiquitous expectation effects on social behaviour and academic performance (Rosenthal and Jacobson, 1968; McDavid and Harari, 1973).

A possible complicating factor in any name stereotype effect may be due to sex roles. While results have not been totally conclusive or consistent it does appear that the sex of subject (evaluator) and sex of target may both be important factors in evaluative judgments. Taynor and Deaux (1975) and Bernard (1979) have shown that females tend to rate performance higher than males, though this was affected by the nature of the task being performed. This finding would be in line with the frequently noted sex difference which portrays females as more person (cf. task) orientated than males (e.g., Douvan and Adelson, 1966). Conversely, evidence also suggests that males are more highly rated in certain abilities and roles (Taynor and Deux, 1973, 1975; Etough and Hughes, 1975; Bernard, 1979). However, evidence from Harris (1975) has shown that college instructors of both sexes rate the performance of males and females equally. Thus, while not conclusive, it seems that in evaluating performance sex roles stereotypes may be an important factor and cannot be divorced from consideration when investigating name stereotype effects.

The experiment reported in this paper examines the influence of the stereotypical value of the name held by both an evaluator and bogus primary school pupils on the grading of essays supposedly written by the pupils. That name stereotypes can influence the marking of essays is now well established (McDavid and Harari, 1973; Harris, 1975). The joint effect of evaluator and pupil name stereotypes remains to be determined. How might the subject's/teacher's name influence attitudes to the performance of a pupil with either an attractive or unattractive name? The investigation of this question constitutes the main aim of the experiment reported here. It was hypothesised that subjects with attractive names would rate a set of essays more highly than subjects with unattractive names. No interaction effect due to the attractiveness of the subject's and pupil's names was expected. It was predicted that both groups of subjects, with attractive and unattractive names, would show equally strong name stereotyping of the pupils' names. In line with these predictions, unnamed essays were expected to be evaluated midway between those labelled with attractive or unattractive names.

Sex of subject and pupil was taken into account though no specific predictions were made due to the ambiguity of the extant literature. At the very least incorporating sex as a subject and target factor served to reduce error within the experiment.

METHOD

Pilot study

The pilot study produced the materials used in the main experiment. The required materials were attractive and unattractive male and female names and a series of six essays of approximate equal standard.

To derive the stimulus names, 25 male and 25 female undergraduates rated 45 female and 35 male christian names on 7-point attractiveness and perceived frequency of occurrence scales. All subjects were British and Caucasian. All stimulus names were English and the lists comprised the names of all the students majoring in psychology at Manchester Polytechnic. The ordering of the attractiveness rating and frequency rating tasks was alternated. No subjects rated their own names.

Mean and standard deviation scores were calculated for the attractiveness and frequency ratings of each name. Four names were selected for use in the main study: these were the male and female names with the highest and lowest mean attractiveness ratings and low standard deviations. The attractive female name was Alison ($\bar{X} = 5{\cdot}97$, sd $= 0{\cdot}87$) and the attractive male name was Steven ($\bar{X} = 4{\cdot}76$, sd $= 1{\cdot}38$). The unattractive female name was Beryl ($\bar{X} = 1{\cdot}32$, sd $= 0{\cdot}52$) and the unattractive male name was Norman ($\bar{X} = 1{\cdot}82$, sd $= 1{\cdot}1$). Previous research has shown a strong correlation between the attractiveness of a name and its frequency of occurrence (McDavid and Harari, 1966). This pilot study also found a strong correlation between the attractiveness and frequency of occurrence of a name ($r = 0{\cdot}64$, $P < 0{\cdot}01$), providing some support for the manner in which attractiveness was assessed.

In order to obtain a set of stimuli essays for the main study, 9- to 10-year-old pupils attending a local junior school were asked to write essays on the theme "I remember, I remember . . .". These were marked by a teacher and typed transcripts of six essays of approximately equal standard were produced.

Main study

Design

A $2 \times 2 \times 2 \times 3$ factorial design was used (sex of subject × attractiveness of subject's name × sex of pupil × attractiveness of pupil's name), with repeated measures on the last two factors. The last factor incorporated a level where no name information was given and only the sex of the pupil was recorded on the essay. This constituted a control condition.

Essays were presented in random order and the names on the essays were counterbalanced to ensure each name appeared on each essay an equal number of times and each essay appeared in each condition an equal number of times.

Subjects

These were 36 male and 36 female British Caucasian undergraduate students. Half the subjects had attractive names and the other half had unattractive names, as determined by the survey of names in the pilot study. Ages were between 18 and 25 years. No subject had been tested in the pilot study or had a name which was used in the experiment.

Procedure

Subjects were tested individually. Subjects were told that the aim of the study was to examine the consistency with which children's essays were marked. They were instructed to read each essay in turn and give it a percentage mark. No guiding criteria were given for the marking. The name (or an indication that this had been lost) or sex of the bogus pupils was recorded at the top right hand corner of each essay. The subject's attention was *not* directed to this personal information. Subjects were thanked and debriefed after marking all six essays.

ANALYSIS AND RESULTS

Analysis was by means of a four-way analysis of variance. Only two effects proved significant. First, there was a significant main effect for the attractiveness of the name of the subject ($F = 4{\cdot}86$, $df = 1, 68$, $P = 0{\cdot}031$). The means for this factor showed that attractively named subjects gave the children's essays higher marks ($\bar{X} = 60{\cdot}66$) than did the unattractively named subjects ($\bar{X} = 56{\cdot}26$).

The second significant effect was due to the bogus pupil's name ($F = 5{\cdot}58$, $df = 2, 136$, $P = 0{\cdot}005$). This factor was subjected to further analysis to investigate the specific hypothesis that the essays of unattractively named pupils would receive lower marks than anonymous pupils who would in turn receive lower marks than attractively named pupils. The subsequent trend test demonstrated a significant linear trend across the means ($\bar{X}$s = 59·4, 60·5 and 60·74, respectively. $F = 9{\cdot}17$, $df = 1, 136$, $P < 0{\cdot}005$).

Both the effects outlined above were as predicted in the initial hypotheses, as was the lack of any significant interaction between the stimulus name and subject name main effects. The lack of an interaction effect suggested that both attractively named and unattractively named subjects held equally strong name stereotypes.

No significant sex differences were noted in the way males and females evaluated the stimulus essays ($F = 0{\cdot}005$, $df = 1, 68$, $P = 0{\cdot}94$) or in the way essays supposedly written by boys or girls were evaluated ($F = 1{\cdot}96$, $df = 1, 68$, $P = 0{\cdot}16$).

DISCUSSION

The results of this study support the initial hypotheses, as outlined in the introduction. In accord with previous research, the attractiveness of the name of the hypothetical pupils was found to exert a significant influence on the marks awarded to essays. Although name stereotype effects may be expected to be pervasive throughout social relationships in general, it is at least encouraging to note that the accentuation/depression of marks on written work can be overcome by ensuring the anonymity of the author.

Somewhat less easy to control is the bias on the subject's essay marking which is due to *his or her own* name stereotype. Of course, there will at least be an equal bias on all the essays marked and, if it is desired to compare groups of essays marked by different individuals, statistical methods can be used to partial out the systematic bias of the markers. It is indeed fortunate that there does not appear to be any interaction effect between the marker's self applied name stereotype and the name stereotype applied to the pupils; this would constitute a far more complex situation and prove far more difficult to control.

The lack of any sex differences in essay marking due either to the sex of the subject or the sex of the pupil was a further encouraging finding of this study, though this may have been due, at least in part, to factors inherent within the experimental design. The subjects used in this study were all relatively young and might possibly be expected to hold less stereotypical views on both sex roles and names than their older counterparts. Although there are teachers of the age of the subjects used in this study, the general teaching population is certainly more varied in age. What little evidence there is does indeed suggest that older, more experienced teachers do indeed show more marked name stereotype effects than their less experienced counterparts (McDavid and Harari, 1973). Boys and girls may enact name stereotypes differentially and this in turn could produce a sex difference effect in a study using an older subject sample (Etough and Hughes, 1975; Bernard, 1979).

To place the above discussion in broader context, it should be recalled that the evidence for sex (of pupil or teacher) as an important factor in the evaluative judgments teachers make of pupils is ambiguous and inconsistent. At the very least consistency is likely to require that subjects in different studies assess written work on a common set of dimensions rather than the broad, variable or frequently unspecified dimensions characteristic of many studies in this area (Taynor and Deaux, 1973, 1975; Etough and Hughes, 1975; Bernard, 1979). A slight variation on this argument suggests that a pupil's ability and achievements are more important in teacher perceptions than gender alone (though this need not deny that gender does have an effect) and that sex biases may reflect real differences in academic performance (Prawat and Jarvis, 1980). The complexities of an ongoing 'real life' classroom situation do, however, make it very difficult to evaluate such general pronouncements.

The preceding discussion has focused on the specific results and design of the experiment reported in this paper, the remaining discussion will focus on more general and theoretical considerations.

A name may be considered a personal characteristic in much the same way as physical appearance, dress, speech style and a variety of other indexical and social-interactional cues. On this basis, many of the principles by which these other factors are governed should also apply to name stereotype effects. And, just as the other factors mentioned above appear to have a reciprocal influence on social relationships, so should name stereotypes. The label constituted by a name would be expected to directly influence its bearer and others who interact with him, and indirectly influence their relationship through mutual expectations and reactions. Names are a type of information exchanged early in most social relationships and as such may represent an important influence on the initial stages of relationship development (Duck, 1977a). This represents a very under-researched aspect of name stereotype effects.

But how do names come to acquire this apparent significance for their bearers and others? Certainly 'mere exposure' would produce some of the effects (Zajonc, 1968) and in conjunction with a 'preference-feedback' mechanism (Colman *et al.*, 1981)—which advocates that exposure to a name will influence preference for it and preference will in turn influence exposure—will go some way towards explaining the *maintenance* of name stereotype effects. This leaves the term 'preference' to be explained. To this end a variety of cognitive and reinforcement models may be invoked (Byrne, 1971; Duck, 1977b), and these will also provide additional bases for the explanation of stereotype effects.

In conclusion, it is argued that name stereotypes constitute one of a network of factors influencing the initiation and development of social relationships. Stereotype effects might thus be expected to be explicable in similar theoretical terms to those used to explain the nature of other antecedents of social relationships. To the end of integrating the literature on name stereotypes with the broader literature on social and personal relationships in general, it is advocated that further research examine the relationship between name stereotypes and the various other antecedents of social relationships, especially the theoretical bases of such relationships.

REFERENCES

BERNARD, M. E. (1979). Does sex role behaviour influence the way teachers evaluate students? *J. educ. Psychol.*, 7, 553-562.

BUCHANAN, B., and BRUNING, J. L. (1971). Connotative meaning of names and nicknames on three dimensions. *J. soc. Psychol.*, 85, 143-144.

BYRNE, D. (1971). The attraction paradigm. New York and London: Academic Press.

COLMAN, A. M., SLUCKIN, W., and HARGREAVES, D. J. (1981). The effect of familiarity on preferences for surnames. *Br. J. Psychol.*, 72, 363-370.

DEAUX, K., and TAYLOR, J. (1973). Evaluation of male and female ability: bias works two ways. *Psychol. Rep.*, 32, 261-262.

DOUVAN, E., and ADELSON, J. (1966). *The Adolescent Experience*. New York: Wiley.

DUCK, S. W. (1977a). *Theory and Practice in Interpersonal Attraction*. New York and London: Academic Press.

DUCK, S. W. (1977b). *The Study of Acquaintance*. Saxon House.

ETOUGH, C., and HUGHES, V. (1975). Teacher evaluations of sex behaviour in children: the role of teacher sex and school setting. *Dev. Psychol.*, 11, 394-395.

GARWOOD, S. G. (1976). First name stereotypes as a factor in self concept and school achievement. *J. educ. Psychol.*, 68, 482-487.

HARRIS, M. B. (1975). Sex role stereotypes and teacher evaluations. *J. educ. Psychol.*, 67, 751-756.

LAWSON, E. D. (1971). Semantic differential analysis of men's first names. *J. Psychol.*, 78, 229-240.

McDavid, J. W., and Harari, H. (1966). Stereotyping of names and popularity in grade school children. *Child Dev.*, 37, 453-459.

McDavid, J. W., and Harari, H. (1973). Name stereotypes and teacher expectations. *J. educ. Psychol.*, 65, 222-225.

Prawat, R. S., and Jarvis, R. (1980). Gender difference as a factor of teachers' perceptions of students. *J. educ. Psychol.*, 72, 743-749.

Rosenthal, R., and Jacobson, L. (1968). *Pygmalion in the Classroom: Teacher Expectation and Pupils Intellectual Development.* New York: Holt, Rinehart and Winston.

Stroebe, W. (1977). Self esteem and interpersonal attraction. In Duck, S. W. (Ed.), *Theory and Practice in Interpersonal Attraction.* New York and London: Academic Press.

Strunk, O. (1958). Attitudes towards one's name and oneself. *J. individ. Psychol.*, 14, 54-67.

Taynor, J., and Deaux, K. (1973). When women are more deserving than men: equity, attribution and perceived sex differences. *J. Person. soc. Psychol.*, 28, 360-367.

Taynor, J., and Deaux, K. (1975). Equity and perceived sex differences: role behaviour as defined by the task, the mode, and the actor. *J. Person. soc. Psychol.*, 32, 381-390.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *J. Person. soc. Psychol.*, 2. Monograph 1-29.

(*Manuscript received 15th September*, 1983)

Br. J. educ. Psychol., 54, 228-234, 1984

BIGGS' SPQ: A BRITISH STUDY OF ITS INTERNAL STRUCTURE

BY M. J. O'NEIL
(*Teesside Polytechnic, Middlesbrough*)
AND D. CHILD
(*School of Education, University of Leeds*)

SUMMARY. An investigation of the study processes of British higher education students using Biggs' Study Process Questionnaire (SPQ) is reported. There is partial confirmation of Biggs' findings that the study process complex, as operationalised by the SPQ, comprises the three orthogonal factor dimensions of utilising, internalising and achieving, where each dimension is made up of a motivational and a strategy subscale. The least satisfactory dimension is " utilising " which appears to be the most unstable scale with few consistent items from study to study. Our research yields evidence to support Biggs' congruence hypothesis that students with particular motivational predispositions tend to adopt a corresponding, linked strategy.

INTRODUCTION

Until recently there was an assumption in formal education that there was a particular collection of good study habits operating across all learning contexts which when adopted would ensure academic success. Recent research has shown this to be groundless (e.g., Entwistle *et al.*, 1971) and the last decade has seen a shift in emphasis towards investigating the relationship between, on the one hand, the student's motivation and conception of his own learning and, on the other, a variety of possible study strategies (Marton and Säljö, 1976; Entwistle and Wilson, 1977; Biggs, 1978). There are, Biggs (1978) contends, several possible ways of achieving success in academic learning: personal preferences, values and motives predispose the student to adopt a distinct study strategy—" *how* we study is a function of *why* we study " (Biggs, 1979a)—*and* congruence between predisposition and strategy leads to success. The earlier prescription of a single collection of ' good study habits ' may therefore be seriously dysfunctional since it may divert the students from the intuitive match they are predisposed to make.

Clarification of such research questions and illumination of precise interrelationships that determine academic performance is offered by Biggs (1978, 1979a, 1979b). He acknowledges, beside predisposition and strategy, the importance of other variables (individual differences in ability, cognitive style and personality, and contexts of learning) and speculates that the study process complex is in fact embedded in a ' presage-process-product ' framework as used, for instance, by models-of-school-learning theorists (e.g., Bloom, 1976; Harnischfeger and Wiley, 1978). Biggs' model (Biggs, 1978, p. 267) is paralleled by that proposed by Svensson (1976, 1979) who describes the study process complex in terms of ' situation, study activities and criterion of learning '.

The Study Process Questionnaire (SPQ) has been constructed by Biggs (1978, 1979a) as a procedure for investigating the study *process* domain of the model in isolation from presage and product variables. Earlier work by Biggs resulted in an 80-item inventory representing an operational definition of those personal qualities which a literature search revealed as significant to academic learning and study. This 80-item questionnaire, subdivided into 10 unidimensional scales, has been shown to have moderate scale internal consistency reliabilities and there has also been strong evidence to support Biggs' contention that the study process domain is composed of three factor dimensions: Reproducing, Internalising and Organising (Biggs, 1978; Watkins and Hattie, 1980).

Subsequent research by Biggs (1979a) has resulted in a re-formulation of the SPQ which now comprises 42 items, still focusing on the three dimensions. Further, he carried out a second-order factor analysis which revealed that each dimension had associated with it a motivational and a strategy aspect. In other words, for a particular dimension students who were motivated in a particular way for studying tended also to adopt a corresponding strategy. Table 1 displays the three dimensions along with corresponding motives and

TABLE 1

RELATIONSHIPS BETWEEN MOTIVE AND STRATEGY FOR BIGGS' THREE DIMENSIONS

Dimension	Motive	Strategy
D_1 *Utilising*	M_1 *Instrumental*: Main purpose is to gain a qualification with pass-only aspirations. This is done as a means of obtaining a better job, more money, or some other extrinsic need.	S_1 *Reproducing*: overall, simply to avoid failure and specifically to focus on minimal content, primarily factual, as prescribed in class handouts, course outlines, etc., and to rote learn this necessary minimum for reproduction in examinations and/or assignments.
D_2 *Internalising*	M_2 *Intrinsic*: Study to actualise interest and competence in particular academic subjects and to develop special interests and abilities; studies are selected therefore that hold maximum intrinsic interest.	S_2 *Meaningful*: read widely and with maximum understanding (independently of course requirements), to integrate various subjects and make them personally meaningful.
D_3 *Achieving*	M_3 *Achievement*: to excel in studies as part of a general competitive approach to life and win high status thereby; more specifically, to study with a view to maximising grades awarded whether or not the material is interesting.	S_3 *Organising*: close orientation to course outlines, follow up all suggested readings, schedule time, behave as " model " student.

(After Biggs, 1979a and 1981.)

strategies which emerged from Biggs' second order factor analyses of the 80-item inventory. In the 42-item version of the SPQ there are seven items to a subscale (see Table 3 for detail).

Whilst Biggs argues that these three dimensions are representative of general tendencies only and that many students exhibit mixed motives and strategies, he nevertheless emphasises that they are predominantly motivated in one of these three ways and use the corresponding, linked strategy.

Additional supportive research work carried out by Hattie and Watkins (1981), using Australian and Filipino university students, has clearly shown for an Australian sample that the SPQ has adequate to good internal consistency coefficients and a factor structure congruent with Biggs' (1978) original findings. For their Filipino sample, however, Hattie and Watkins report low to moderate internal consistency coefficients and, unfortunately, factor analyses which do not support Biggs' motive-strategy hypothesis.

To summarise the findings so far, Biggs (1978) used Canadian and Australian university students' responses to the SPQ to derive a ' virtually identical ' factor solution. Hattie and Watkins (1981) found congruence between the Biggs' model and their Australian sample, but incongruence with their Filipino data. Clearly, Hattie and Watkins are justified in their request for more evidence to test the extent to which SPQ is culturally or country bound and additionally for further testing of the validity of SPQ as a measure of study motives and behaviour.

This paper, then, reports research using SPQ (42-item version) in a British polytechnic setting. The primary aims of this research were to investigate:

(i) the internal consistency of the SPQ scales using Cronbach's α-coefficient;
(ii) the factor dimensions of the SPQ items and scales;
(iii) the comparability between this factor solution and that of Biggs;
(iv) oblique solution and a second order analysis;
(v) any relationship between students' preferred motives and strategies.

METHOD

Sample

The subjects were 277 full- and part-time CNAA undergraduates and advanced diploma students (age range 18-27 years with a mean age of 20·5 years) at an English polytechnic—a small, technology-oriented institution in a northern urban/industrial centre. Missing values for some responses to some questions reduced the final sample size to 245 (169 males and 76 females).

Procedure

Heads of departments in the polytechnic were asked to assist in the research project and to invite returns from 30 or so of their students in several classes. In the event, nine departments (out of 14) participated, yielding a final sample of students with a broad disciplinary composition reflecting the technology orientation of the polytechnic's work. Students were tested in class groups on a voluntary return basis and the response rate was 85 per cent. All the testing took place in 1981.

RESULTS

(i) *Reliability of the scales*

The reliabilities of the scales and subscales of the SPQ Form UC were determined using Cronbach's alpha coefficient. Table 2 shows these and the values are moderate. As will become apparent later, there is less confidence in the utilising scale, particularly the subscale for strategy, than the others.

(ii) *Factor dimensions of SPQ items*

The correlation between the 42 items for the 245 complete questionnaires gave a positive manifold and an average value of +0·15. Factor analysts are still debating the question of how large a sample should be for a given number of variables. The suggested ratio of sample to variables ranges from one (Nunnally, 1978) through at least five (Gorsuch, 1974) to ten (Aleamoni, 1976). Baggeley (1982) gives some useful guidance on this matter. He derives a formula using the mean absolute value of the correlations between the variables. For this study, with the above value of 0·15 and 42 variables, the corresponding minimum number of subjects is 168. With a sample of 245 the number is well over this critical value.

The original solution by Biggs employed the varimax method. This was repeated in the present case using the SPSS package (Nie *et al.*, 1975). A principal factor analysis using iterations to obtain an estimate of the communalities, orthogonal rotation using the varimax solution for principal factors having eigen values greater than one, oblique rotation using the direct-oblimin method and a second-order factor analysis of this direct-oblimin solution were undertaken. Thirteen factors had eigen values greater than one and these were rotated. Of these, only six were clearly interpretable in both the varimax and oblimin solutions using 0·25 as the minimum level for a salient loading.

TABLE 2

RELIABILITY OF SPQ SCALES AND SUB-SCALES USING CRONBACH'S ALPHA COEFFICIENT (N = 245)

	Scales	α-coefficient
D_1	Utilising	0·69
D_2	Internalising	0·76
D_3	Achieving	0·78
	Sub-scales	
M_1	Utilising motive	0·53
S_1	Utilising strategy	0·56
M_2	Internalising motive	0·64
S_2	Internalising strategy	0·65
M_3	Achieving motive	0·72
S_3	Achieving strategy	0·73

TABLE 3

SALIENT FACTOR LOADINGS FOR BIGGS' (VARIMAX) TWO SAMPLES AND PRESENT STUDY (VARIMAX AND DIRECT OBLIMIN)

Sub-scale	Question no.	Salient loadings																							
		Biggs' Varimax (univs.)						Biggs' Varimax (CAEs)						Present Varimax						Present Oblimin					
		1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
M_1	1	655						671						496						460					
	7																								
	13	608				347		607				290		437				337		427					
	19																								
	25		594							−337															
	31	392		−476						−477				508						495					
	37	687						686						559						568					
S_1	4		678						289						606						585				
	10																								
	16																								
	22		557				−329								567					248	479				
	28		268						514					342											
	34								723					261	248		−476						−439		
	40								714					401						323					
M_2	2			578						598	252					634						614			
	8									304	268					286	392					275	289		
	14			537						447						400						364			
	20			597						599						625						608			
	26			556	324					598						549						539			
	32															276	372						387		
	38	272																							
S_2	5				420						650						477						372		
	11				683					352	510														
	17			246	261		283			268	385		325				404						387		
	23				720						708						480						255		
	29		−304				455			347			515												
	35		−377				355						520			310									
	41				701						650						473						301		
M_3	3					724						707						455						311	
	9					649	247					618						481	413					487	315
	15					643						699						585						560	
	21									366		321						376						307	
	27											494													
	33					438						706						572						507	
	39											399						327						240	
S_3	6						540						573												
	12						656						474						683						673
	18						676						539						625						599
	24						660						677						536						461
	30				315		506						559						270						
	36						649						704												
	42						572						447						587						589

The factors (sub-scales) are:
1 and M_1 Utilising motive; 2 and S_1 Utilising strategy; 3 and M_2 Internalising motive;
4 and S_2 Internalising strategy; 5 and M_3 Achieving motive; 6 and S_3 Achieving strategy.
Decimal points omitted.
Salience $\geq$ 0·25.
Biggs' Varimax (univ.) = University sample of 827.
Biggs' Varimax (CAEs) = College of Advanced Education, N = 1,638.

The salient loadings shown in Table 3 from the factor pattern matrices give at least a *prima facie* case for believing there to be similarities. Subscales M_2, M_3, S_2 and S_3 look particularly robust when one compares the varimax patterns. Another interesting observation is that, wherever a significant loading appears outside the block of loadings associated with a subscale, it is either of the same dimension (M_1S_1, M_2S_2, M_3S_3), or is one of the same motives or strategies as the subscale.

(iii) *Comparability of analyses*

There are several methods for comparing factors from two or more studies using the same variables but different subjects. A useful method has been worked out by Cattell and Baggeley (1960; see also Cattell *et al.*, 1969, for modifications). There is a helpful chapter on this subject in Gorsuch, Chapter 13, 1974. It is particularly appropriate where markers can be assumed which in this case are the subscale items of Biggs' analysis. The method also avoids capitalising on chance.

TABLE 4

TESTING THE SIMILARITY BETWEEN FACTORS FROM BIGGS AND THIS STUDY USING *s* THE SALIENT VARIABLE SIMILARITY INDEX

	No. of salient loadings corresponding between the present study and:			
	Biggs (univ)		Biggs (CAE)	
Sub-scale	No.	Sig.	No.	Sig.
M_1	4	very high	3	0·01
S_1	2	NS	2	NS
M_2	4	very high	5	very high
S_2	4	very high	4	very high
M_3	4	very high	6	very high
S_3	5	very high	5	very high

Only those variables within the corresponding blocks of the sub-scales have been considered here.

Clearly, from Table 4, five of the six subscales compare most favourably. S_1, the utilising strategy, appears to be the most unstable and has few consistent items from study to study. The internalising and achieving dimensions are very sound and the latter replicates the findings of many studies relating to achievement motivation (e.g., McClelland *et al.*, 1976; Brookover *et al.*, 1967; Marton and Säljö, 1976; Entwistle *et al.*, 1979).

(iv) *Oblique solution and second order factor analysis*

This was done to see if a clearer solution emerged and to test the orthogonality of the scales. Table 3 includes the direct oblimin factors and they do not add anything to the clarity of the varimax solution. Perhaps more significant is the inter-correlation of these factors. These values are all low and would lead us to conclude that the six subscales are virtually independent.

The second order analysis using the correlation matrix from the first order oblimin solution gave no discernible pattern. This ties in with the near independence of the subscales alluded to above.

(v) *Motive-strategy relationships*

In order to test Biggs' (1978, 1981) congruence hypothesis that students who are motivated in a particular way tend to adopt a corresponding, linked strategy, we looked at the inter-correlations among the actual scores obtained by the sample on the six subscales. From the data in Table 5 there is strong evidence to support Biggs' claim for motive-strategy associations in the same dimension (M_1-S_1, M_2-S_2 and M_3-S_3 correlations have high levels of significance).

TABLE 5

CONGRUENT AND INCONGRUENT MOTIVE-STRATEGY CORRELATIONS FOR BIGGS' TWO SAMPLES AND PRESENT STUDY

Motive-strategy	Biggs' univ.	Biggs' CAEs	Present study
M_1-S_1	43	41	49**
M_1-S_2	−10	02	−06
M_1-S_3	10	11	13
M_2-S_1	−17	−11	−11
M_2-S_2	62	58	50**
M_2-S_3	31	35	34**
M_3-S_1	22	20	36**
M_3-S_2	24	32	21*
M_3-S_3	31	31	41**
M_1-M_2			−08
M_1-M_3			40**
M_2-M_3			19*
S_1-S_2			−19*
S_1-S_3			14
S_2-S_3			40**

Decimal points omitted.
* $P < 0{\cdot}01$; ** $P < 0{\cdot}001$.

Furthermore, as Table 5 reveals, there is high point-to-point correspondence between the present findings and Biggs' (1981) data.

DISCUSSION

This replication study of the internal structure of the SPQ using a British sample lends general support to findings reported by Biggs (1978, 1979a, 1981) and Hattie and Watkins (1981). The reliabilities using Cronbach's alpha coefficient for the scales and subscales were favourable; all but the S_1 subscale (utilising strategy) showed high similarity to Biggs' original solution (Table 4); the oblique solution gave small correlations between the identifiable factors, showing independence of the subscales; the correlations between corresponding motives and strategies (M_1-S_1 etc.) were highly significant and supportive of Biggs' contention that there should be congruence between the professed motives of students and their corresponding strategies of study.

However, there are some anomalous features which require attention. The utilising dimension has only moderate internal consistency (α-coefficient). The number of items in the analysis shown in Table 3 (3 or 4 across all the studies reported) is marginal in defining clearly the utilising subscales proposed by Biggs. There are some items which do not have a salient value in any of the six identifiable factors extracted in any solution (items 7, 10, 16, 19). The only subscale which could not be located in both solutions (Table 4) was S_1.

With these doubts about the soundness of the utilising dimension in mind, particularly S_1 (utilising strategy), the following tentative observations are made about the general pattern of other highly significant interactions. Both M_1 and S_1 are significantly correlated with M_3. A closer look at this suggests that it may be due to a 'fear of failure' overlap (*vide* McClelland *et al.*, 1976). M_3 is also linked with S_2. The negative correlation between S_1 and S_2 is in line with Biggs' hypothesis that students who are oriented towards a 'meaningful' study strategy are least likely to use a 'reproducing' one, and *vice versa.* All these connections, and the other significant intercorrelations among motives and strategies, would suggest that there is a major split between the dimensions (D_1 versus D_2 with D_3) with M_3, however, having a foot in both camps. This outcome is reminiscent of Marton's (1976) distinction between 'surface level' and 'deep level' processing of material to be learned. Indeed, Biggs (1983) provides this information in an, as yet, unpublished form, giving normative data for the six subscales, when he lists the 'deep' score as the sum of M_2, M_3, S_2 and S_3, and the 'surface' score as the total of M_1 with M_3 and S_1 (*vide* Biggs and Kirby, 1983).

Two obvious lines of research are needed. One is to explore the robustness of the utilising dimension. The other is to look at the actual performance of students, both in study and in examinations, to test the validity of their responses about preferred study strategies and to discover whether congruence-incongruence of motive and strategy has any bearing on performance.

ACKNOWLEDGMENTS.—We would like to thank the staff and students of the polytechnic for their help and Professor John Biggs for letting us have details of his analysis and various unpublished papers, ideas from which have been incorporated in this paper.

REFERENCES

ALEAMONI, L. M. (1976). The relation of sample size to the number of variables in using factor analysis techniques. *Educ. psychol. Meas.*, 36, 879-883.

BAGGELEY, A. R. (1982). Deciding on the ratio of number of subjects to number of variables in factor analysis. *Mult. exp. clin. Res.*, 6, 81-85.

BIGGS, J. B. (1978). Individual and group differences in study processes. *Br. J. educ. Psychol.*, 48, 266-279.

BIGGS, J. B. (1979a). A Survey of Students' Study Processes. Unpublished manuscript, University of Newcastle, NSW, Australia (personal communication).

BIGGS, J. B. (1979b). Individual differences in study processes and the quality of learning outcomes. *Higher Educ.*, 8, 381-394.

BIGGS, J. B. (1981). Motivational patterns, learning strategies and subjective estimates of success in secondary and tertiary students. Paper presented to Annual Conference of Australian Association for Research in Education, Adelaide, November.

BIGGS, J. B. (1983). SPQ Norms. Unpublished data sheet, University of Newcastle, NSW, Australia (personal communication).

BIGGS, J. B., and KIRBY, J. R. (1983). Approaches to learning in universities and CAEs. *Vestes*, 27, 3-9.

BLOOM, B. S. (1976). *Human Characteristics and School Learning*. New York: McGraw-Hill.

BROOKOVER, W. B., ERIKSON, E. L., and JOINER, L. M. (1967). *Self-concept of Ability and School Achievement*—III, Final Report on Co-operative Research Project No. 2831. East Lansing, Michigan State University.

CATTELL, R. B. and BAGGELEY, A. R. (1960). The salient variable similarity index for factor matching. *Br. J. stat. Psychol.*, 13, 33-46.

CATTELL, R. B., BALCAR, K. R., HORN, J. L., and NESSELROADE, J. R. (1969). Factor matching procedures: an improvement of the index; with tables. *Educ. psychol. Meas.*, 29, 781-792.

ENTWISTLE, N. J., HANLEY, M., and HOUNSELL, D. J. (1979). Identifying distinctive approaches to studying. *Higher Educ.*, 8, 365-380.

ENTWISTLE, N. J., NISBET, J., ENTWISTLE, D., and COWELL, M. D. (1971). The academic performance of students. I. Prediction from scales of motivation and study methods. *Br. J. educ. Psychol.*, 41, 258-267.

ENTWISTLE, N. J., and WILSON, J. (1977). *Degrees of Excellence: The Academic Achievement Game*. London: Hodder and Stoughton.

GORSUCH, R. (1974). *Factor Analysis*. Philadelphia: Saunders.

HARNISCHFEGER, A., and WILEY, D. E. (1978). Conceptual issues in models of school learning. *J. Curric. Stud.*, 10, 215-231.

HATTIE, J., and WATKINS, D. (1981). Australian and Filipino investigations of the internal structure of Biggs' new Study Process Questionnaire. *Br. J. educ. Psychol.*, 51, 241-244.

MARTON, F. (1976). On non-verbatim learning IV—theoretical and methodological considerations. *Scand. J. Psychol.*, 17, 125-128.

MARTON, F., and SÄLJÖ, R. (1976). On qualitative differences in learning—II. Outcome as a function of the learner's conception of the task. *Br. J. educ. Psychol.*, 46, 115-127.

MCCLELLAND, D. C., ATKINSON, J. W., CLARK, R. A., and LOWELL, E. L. (1976). *The Achievement Motive*. New York: Irvington.

NIE, N. H., HULL, C. H., JENKINS, J. G., STEINBRENNER, K., and BENT, D. H. (1975). *Statistical Package for the Social Sciences*. (2nd ed.) New York: McGraw-Hill.

NUNNALLY, J. C. (1978). *Psychometric Theory* (2nd ed.). New York: McGraw-Hill.

SVENSSON, L. (1976). *Study Skill and Learning*. Acta Universitatis Gothoburgensis, Gothenburg.

SVENSSON, L. (1979). On qualitative differences in learning III.—Study skill and learning. *Br. J. educ. Psychol.*, 47, 233-243.

WATKINS, D., and HATTIE, J. (1980). An investigation of the internal structure of the Biggs Study Process Questionnaire. *Educ. psychol. Meas.*, 40, 1125-1130.

(*Manuscript received 27th September*, 1983)

Br. J. educ. Psychol., 54, 235-238, 1984

THE RELATIONSHIP BETWEEN LOCUS OF CONTROL AND TEACHER BURNOUT

By T. C. McINTYRE

(Department of Special Education, Eastern Illinois University)

SUMMARY. The purpose of this research was to investigate the relationship between locus of control, as measured by the Nowicki-Strickland Adult Scale, and burnout as measured by the Maslach Burnout Inventory (MBI). A survey form containing these instruments was mailed to all special education teachers in 28 school districts in the north-eastern United States. Of the initial 684 mailings, 469 were returned (69 per cent). An analysis of the data using Pearson product-moment correlations revealed significant correlations between increasingly external locus of control orientation and increased perceptions of burnout on four of the six subscales of the MBI. Findings are discussed and a suggestion for future research presented.

INTRODUCTION

Stress and burnout among teachers has, in the last few years, become a popular topic in the professional literature and at conferences of educational organisations. To date, researchers have identified numerous sources of stress (Bensky *et al.*, 1980; Fimian and Santoro, 1983; Schwab, 1983) and personal background characteristics of educators who appear to be experiencing burnout to a greater degree than their colleagues (Schwab, 1980; Crane, 1981; McIntyre, 1982; Presley and Morgan, 1982). However, the relationship of one's personality to degree of perceived burnout has been only superficially investigated. The present study investigated the relationship between one personality variable, locus of control, and experienced burnout.

Locus of control is concerned with " the question of whether or not an individual believes that his own behavior, skills or internal dispositions determine what reinforcements he receives " (Rotter *et al.*, 1972). Individuals are represented along a continuum, with " internals " feeling that they are effective in controlling their destiny and determining the occurrence of reinforcement, and " externals " believing that forces beyond their control (e.g., fate, chance, luck, powerful others) determine the occurrence of specified events. It can be described as a self appraisal of the degree to which an individual views him/herself as having a causal role in determining specific events, or a series of situations which are perceived as being similar.

One's ability to cope with environmental stress appears to be influenced by one's locus of control. Persons with an internal locus of control appear to handle environmental stress better than persons with an external orientation. For instance, in comparison with internals, those with a more external orientation report more debilitating anxiety, more neurotic symptoms, and more self-punitiveness in response to frustration (Efran, 1971; Butterfield, 1964; Rotter, 1966; Tolar and Rezinkoff, 1967; Watson, 1967; Feather, 1968; Platt and Eisenman, 1968; Goss and Morisko, 1970; Hountras and Scharf, 1970).

Several investigations have researched locus of control with respect to teachers (Carter, 1969; Monroe, 1969; Benoff, 1976; Healey, 1976; Ferguson, 1980; Kremer and Ben, 1980; Scheck and Rhodes, 1980), and one in particular has looked at the relationship between locus of control and teacher stress among regular educators.

Using a questionnaire format, Kyriacou and Sutcliffe (1979) assessed the relationship between locus of control and reported stress. The teachers were asked to rate the question, " In general, how stressful do you find being a teacher?" on a five-point scale from " not stressful " to " extremely stressful." Also included in the questionnaire was a Likert scale version of the Rotter I-E instrument to measure locus of control. The correlation between self reported stress and a belief in external control was found to be positive and significant.

A study in a similar vein to the locus of control construct, conducted by Meadow (1981), compared answers to the question, " Are you given the opportunity to influence your own work situation or assignment?" with responses to the subscales on the Maslach Burnout Inventory. Results showed that those reporting a greater degree of influence in their work environment also reported lesser emotional exhaustion, lesser depersonalisation, and greater personal accomplishment than those who reported lesser influence. As degree of influence scores lowered, perceived feelings of burnout increased. The present study sought to investigate further the connection between locus of control and burnout.

METHOD

Population

The population included all resource room and self-contained room special education teachers from 28 public school districts in Connecticut and Massachusetts (N = 684). Other school personnel were not included in the analysis of data.

Instrumentation

Burnout was measured by the Maslach Burnout Inventory (Maslach and Jackson, 1981). The inventory consists of three normative scales, each of which measure an aspect of burnout. The emotional exhaustion subscale assesses the degree to which persons feel emotionally drained by their jobs. The depersonalisation subscale determines the extent to which individuals view themselves as having negative attitudes toward their clients and co-workers. Lastly, the personal accomplishment subscale assesses the degree to which respondents feel a sense of professional gratification from their occupation. Each of these subscales is comprised of two dimensions; frequency of feelings, and intensity of feelings.

Locus of control was assessed by the Adult Nowicki-Strickland Internal-External Control Scale (ANS-IE) (Nowicki and Strickland, 1973). The ANS-IE scale consists of 40 yes-no questions which assess the degree to which a person perceives a causal relationship existing between their behaviour and resultant reinforcement.

Data analysis procedures

The relationship between locus of control and teacher burnout was assessed by the use of a Pearson Product-Moment Simple Correlation technique. An 0·05 level of significance on the LSD-MOD post-test procedure was used.

RESULTS

Of the 684 special education teachers who received the survey, 469 returned the questionnaire. This represented a response rate of 69 per cent. The elimination of unusable instruments and the deletion of all cases with missing values left 399 surveys available for analysis. Locus of control was found to correlate significantly with four of the six burnout dimensions, and a trend was found for a fifth dimension. Significant positive correlations were found between locus of control scores and the frequency of feelings of emotional exhaustion, and intensity of feelings of depersonalisation. That is, as locus of control scores increased, indicating a more external orientation, teachers felt more emotionally exhausted and became more depersonalised toward their students.

A negative correlation was found between locus of control and frequency of feelings of personal accomplishment. Therefore, as a person scores in a more external direction, they experience fewer feelings of personal accomplishment. A trend ($P = 0{\cdot}062$) was found for strength of feelings of personal accomplishment. (This trend approached nearer to significance ($P = 0{\cdot}053$) when the effects of age, teaching position, and level of attained education were statistically controlled by a partial correlation procedure.) Table 1 presents these results.

These results indicate that as teachers report a more external orientation they also report a greater degree of burnout. These findings are in concert with what would be expected from reviewing related literature which shows internals coping better with stress. The findings are also similar to those of Kyriacou and Sutcliffe (1979) who found teachers with an external orientation reporting significantly more job related stress. Meadow's (1981) results would also be in concert with the findings of the present study.

TABLE 1

PEARSON CORRELATION COEFFICIENTS BETWEEN BURNOUT SCORES AND LOCUS OF CONTROL

N = 399

Dependent variables	Locus of control
Emotional exhaustion (Frequency)	0·1652*
Emotional exhaustion (Intensity)	0·1693*
Depersonalisation (Frequency)	0·0794
Depersonalisation (Intensity)	0·1148*
Personal accomplishment (Frequency)	0·1103*
Personal accomplishment (Intensity)	0·0919

* $P < 0·05$

DISCUSSION

This study reveals that a personality variable, locus of control, is related to certain aspects of burnout. A question remains, however, as to whether one variable might be causing the other to change. That is, does the process of " burning out " change one's locus of control, or does one's original orientation make one more susceptible to burnout? One strategy, given research which indicates that locus of control can be changed (Gottesfeld and Dozier, 1966; Harrow and Terrante, 1969; Smith, 1982), would be to attempt change in orientation for external, burned out teachers. This change could be verified by readministration of the locus instrument, and reassessment of perceptions of burnout via readministration of the MBI in order to see if change in locus of control effected levels of burnout.

Another option would be to conduct a study involving internally oriented persons entering the teaching profession. Periodic retesting on the MBI and locus instrument would document the connection, if any, between the two variables.

Further examination and investigation into the connection between locus of control or other personality variables and burnout is warranted and necessary.

REFERENCES

BENOFF, M. (1976). Consultation in urban and suburban elementary schools: teachers' expectations and interactions as a function of internal-external locus-of-control. DAI Vol. 37 (04), Sec. A, p. 2068.

BENSKY, J. M., SHAW, S. F., GOUSE, A. S., BATES, H., DIXON, B., and BEANE, W. E. (1980). Public Law 94-142 and stress: a problem for educators. *Except. Children*, 47, 24-29.

BUTTERFIELD, E. C. (1964). Locus of control, test anxiety, reaction to frustration, and achievement attitudes. *J. Person.*, 32, 298-311.

CARTER, R. M. (1969). Locus of control and teacher expectancy as related to achievement of young school children. *Dissert. Abstr. Internat.*, 30 (11), Sec. A, 4674.

CRANE, S. (1981). An analysis of the relationship among personal and professional variables, role ambiguity, role conflict, and perceived burnout of special education teachers. Doctoral dissertation, University of Connecticut.

EFRAN, J. (1971). Some personality determinants of memory for success and failure. *Psychol. Rep.*, 28, 619-640.

FEATHER, N. T. (1968). Change in confidence following success or failure as a predictor of subsequent performance. *J. consult. clin. Psychol.*, 32, 110-119.

FERGUSON, T. (1980). An investigation of various relationships among teachers on a measure of self-concept and locus of control with respect to age, sex, and selected demographic variables. Doctoral dissertation, University of Connecticut.

FIMIAN, M. J., and SANTORO, T. M. (1983). Sources and manifestations of occupational stress as reported by full-time special education teachers. *Except. Child.*, 49, 6, 540-543.

GOSS, A., and MORISKO, T. E. (1970). Relation between a dimension of internal-external control and the MMPI with an alcoholic population. *J. consult. clin. Psychol.*, 34, 189-192.

GOTTESFELD, H., and DOZIER, G. (1966). Changes in feelings of powerlessness in a community action program. *Psychol. Rep.*, 19, 978.

HARROW, M., and TERRANTE, A. (1969). Locus of control in psychiatric patients. *J. consult. clin. Psychol.*, 33, 582-589.

HEALEY, G. R. (1976). The relationship between locus of control and location of school of acceptance of behavior by elementary school teachers. Doctoral dissertation, University of Connecticut.

HOUNTRAS, P. T., and SCHARF, M. C. (1970). Manifest anxiety and locus of control of low achieving college males. *J. Psychol.*, 74, 95-100.

KREMER, L., and BEN, P. M. (1980). Teachers' characteristics and their reflection in curriculum implementation. *Stud. educ. Evaluation*, 6, 73-82.

KYRIACOU, C., and SUTCLIFFE, J. (1979). A note on teacher stress and locus of control. *J. occup. Psychol.*, 52, 227-228.

MASLACH, C., and JACKSON, S. E. (1981). *The Maslach Burnout Inventory.* Palo Alto, Calif.: Consulting Psychologists Press.

MCINTYRE, T. (1982). An investigation of the relationships among burnout, locus of control, and selected personal/professional variables in special education teachers. (Doctoral dissertation, University of Connecticut, 1981.) *University Microfilms International*, 1982.

MEADOW, K. P. (1981). Burnout in professionals working with deaf children. *Am. Annals Deaf*, 2, 13-22.

MONROE, D. S. (1969). An investigation of the effects on job satisfaction of the interaction of teacher's orientation of locus of control and the degree of bureaucratisation of selected secondary schools. *Dissert. Abstr. Internat.*, 30 (7), Sec. A, 2769.

NOWICKI, S., and STRICKLAND, B. R. (1973). A locus of control scale for adults. *J. consult. clin. Psychol.*, 40, 148-154.

PLATT, J. J., and EISENMAN, R. (1968). Internal-external control of reinforcement, time perspective, adjustment, and anxiety. *J. gen. Psychol.*, 79, 121-128.

PRESLEY, P., and MORGAN, H. (1982). Teacher burnout in special education—myth or reality? *Illinois Council for Exceptional Children Quarterly*, 31 (2), 19-22.

ROTTER, J. B. (1966). Generalised expectancies for internal versus external control of reinforcement. *Psychol. Monog.*, 80 (1), 1-28.

ROTTER, J. B., CHANCE, J., and PHARES, E. J. (Eds) (1972). *Applications of a Social Learning Theory of Personality.* New York: Holt, Rinehart and Winston.

SCHECK, D. C., and RHODES, G. A. (1980). *The Relationship between Junior High School Teachers' Rated Competence and Locus of Control.* ERIC Document Reproduction Service No. ED224221.

SCHWAB, R. L. (1980). An analysis of the relationship between role conflict, role ambiguity, personal variables, and perceived burnout among teachers. Doctoral dissertation, University of Connecticut.

SCHWAB, R. L. (1983). Teacher burnout: moving beyond " psychobabble ". *Theory Into Practice*, 22 (1), 21-26.

SMITH, J. (1982). A study of the relationship of flexibility to self-concept and perception of locus of control among elementary school teachers. Doctoral dissertation, University of Connecticut.

TOLAR, A., and RESNIKOFF, M. (1967). Relation between insight, repression-sensitisation, internal-external control, and death anxiety. *J. abnorm. Psychol.*, 71, 426-430.

WATSON, D. (1967). Relationship between locus of control and anxiety. *J. person. soc. Psychol.*, 6, 593-597.

(*Manuscript received* 10*th October*, 1983)

BOOK REVIEWS

BORMAN, K. M. (Ed.) (1982). *The Social Life of Children in a Changing Society.* London: Lawrence Erlbaum, N.J., pp. v+294, c. £19·95.

Generally speaking, I would want to read any book which contained writing by Norman Denzin, and this volume, edited by and with substantial contributions from Kathryn Borman, proved no exception. Among her undoubted skills of scholarship and attention to detail are those of skilful synthesis and clarity of expression (*vide* ' Cognition and Culture: two perspectives '). But there are blemishes in the book as a whole. Its organisation, which Dr. Borman makes abundantly clear in the introduction, does in reality seem less well articulated, and the levels of contributions themselves vary considerably. Moreover, the volume, despite references *en passant* to English and European work, for the most part is markedly North American and possessed of all those unreadable qualities common to many such social science perspectives. It is not free of jargon and one or two papers, notably that by Sutton-Smith on a performance theory of peer relations, badly need ' unpacking ' and simplifying. It is not that they are not worth reading, more that they are too dense and written, one feels, for an in-group of the cognoscenti, rather than for general information.

Dr. Borman tells us that the book resulted from an interdisciplinary symposium on childhood socialisation and that the major purpose of that symposium was to provide an integrative perspective of the social development of young children. This means that the three major disciplines (anthropology, sociology and psychology) are represented together with work by people who have, in the words of the editor, ' self-consciously adopted an interdisciplinary framework '. Kathryn Borman maintains that four propositions guided her organisation of the book into its four sections. These were: that the socialisation of children is critically dependent on certain significant others (parents, peers); that language is a key factor; that delay through physical and social handicap needs special consideration; that cross-cultural perspectives provide insight into important variation. Few would quarrel with her premises or her plan, particularly since she and Fishbein contribute a masterly and easily-read overview as introduction, and Gideonse a sensitive and timely conclusion on the politics of childhood.

Denzin's notes toward a phenomenology of childhood *are* indeed notes; tantalisingly brief, but highly readable for all that and they take one right to the stuff of childhood. For this reader Denzin weaves little spells. His language is clear, simple, often elegant; he uses images caught from biography which are both apt and poignant. A. A. Milne's passage, on the toad as a symbolic world for the boys and later the account of Jeffrey dying of leukemia and knowing it; moreover, actually, as Denzin puts it, ' orchestrating ' the accompaniment to his death; all this is fitted into an analysis which takes child self-awareness and meaning seriously, which shows convincingly that children do have ' access to their own histories ' and are knowing participants in the construction of their own reality by the age of three or so.

There is plenty on language and communication in this book—as one might expect, most of it in that heavy and overblown portentous prose which one associates with expert observers on communication. Why can't people interested in communication actually communicate that interest? Perhaps, as a colleague once drily remarked to me, ' academics tend to profess that which they do worse '! Certainly the paper by Gearhart and Hall on ' Internal State Words ' does have important things to say on coding—but, for this reader, decoding *their* description was tough going. Ogbu's paper on a culture-ecological approach to socialisation was quite a find, however. Short it may be—but a good starting point for further thought and research.

Books, perhaps, like other forms of communication, do—or no not—' speak to one's condition, ' as the Quaker phrase runs. This book has lots of goodies within it. For me the ' significant others ' section and the cross-cultural perspective spoke much to interest and current context. Certainly Kathryn Borman has done a scholarly and careful job. But I wonder who will read the *whole* book?

PHILIP GAMMAGE

MARTLEW, M., (Ed.) (1983). *The Psychology of Written Language: Developmental and Educational Perspectives.* Chichester: John Wiley & Sons, pp. vi+428, c. £24·75.

There are some good papers here. Cooper and Matsuhashi, in ' A Theory of the Writing Process ', ask what decisions a writer makes and give sensible answers. The paper is of great interest particularly on sentence roles and production, and this section could well have been expanded at the expense of the unduly dutiful survey of ' authorities '.

Another important chapter is ' The Development of Evaluative, Diagnostic and Remedial Capabilities in Children's Composing ' by Scardamalia and Bereiter. The writers offer a choice of prompts to children to guide their revisions. The revised versions are no improvement overall, perhaps because revising threatens the ' hold on the composing process as a whole ' (p. 93). Nevertheless the list of prompts, because they are approximate to monitoring processes, may be helpful pedagogically when internalised, amongst other classroom techniques. Revision is the fashionable ' answer ' in US at present, and needed such a cool look.

Martlew's piece, ' Problems and Difficulties; Cognitive and Communicative Aspects of Writing ', also studies the process by examining the time intervals as teenagers compose; and her detailed discussion of specific cases is valuable. Papers by two young hopefuls called Vygotsky and Luria aren't at all that bad though one has to search hard for the sources. There are authoritative chapters on for instance Orthography (Read), Handwriting (Thomassen and Teulings) and ' Writing Disorders Associated with Focal Cortical Lesions ' (Marcie). Poole writes on ' Socioeconomic Status and Written Language '. There is a scrupulous piece by Bryant and Bradley, ' Psychological Strategies and the Development of Reading and Writing ', which is, however, about reading: its inclusion is scarcely justified.

The mere listing of such titles in fact throws up a problem with the collection. There is no rationale offered for it. Nothing is said about the state of the art, and thus the context of the papers; there is no attempt to evaluate them, or to justify the range. For this the general editor must bear (policy) responsibility. It is not good enough to bundle papers together without interpretation, particularly at £24·75, however nice the silver covers.

The study of writing has moved on from the impetus given it by social and educational linguistics into the field of psychology, and the work of the cognitive psychologists, represented outstandingly here by Bereiter and Scardamalia, gives leadership. However, nothing is included of the equally notable impulse from developmental psychology represented by Odell, Harrison, or the Crediton Project. Again work is going on, not only about process, but also on produce, and evaluation. A rationale would have given us the principles of selection. Some justification for the absence of these now very visible topics is that the book is a reprint of the 1981 American edition, though we are not told this.

Such a rationale might also have given us a reason for including the papers by Hunt and by Watson, which continue the transatlantic obsession with counting trivial features of language in terms of t-units and ' syntatic maturity ' produced by sentence combining. Hunt's paper has its own self-destruct mechanisms. When he says he does not know the relationship of ' syntatic maturity ' (not defined) to writing quality (p. 123) one may well ask if this approach can yield the answer; otherwise what is the point of it. In fact such research has little to do with writing and much to do with class control and marketing; combining sentences is a time-consuming easily-marked exercise and is thus immensely popular. If something were sought about ' count ' measures it could have been from Harpin's SSRC report 757 which quietly sees off (once again!) the sillier parts of the business.

ANDREW WILKINSON

MILES, T. R. (1983). *Dyslexia: The Pattern of Difficulties.* St. Albans: Granada, pp. 225, p. £9·95.

Professor Miles' latest book is aimed at both the interested layman and the specialist. Consequently it is necessary to examine its utility to both these groups of readers.

It is, in many ways, an admirable book for the layman. Written in the first person in a clear and uncluttered style it gives an indication of the author's long dedication to the study of dyslexia, and of the help he has given countless sufferers. The book commences with a succinct history of the development of the notions of dyslexia from the pioneering work by Hinshelwood, Pringle Morgan and Orton through the work of McDonald Critchley and the Bullock and Tizard Reports to recent developments in cognitive psychology. The ensuing chapters describe in detail the assessment procedures which have been adopted by the Bangor team, with indications of the difficulties which may be encountered in the employment of verbally loaded instruments, and the dangers of interpreting psychometric scores with any degree of precision. There is then a discussion of the more subtle signs which the author's experience has prompted him to look for in clinical interviews.

The following chapter contains what are rapidly becoming *de rigeur* in books on dyslexia —the case histories. Their inclusion certainly makes interesting reading, and illustrates selected cases of the condition, though it is to be hoped that the layman will bear in mind the preceding discussion of variability in presenting symptoms, a feature which again receives prominence in the following chapters.

From chapter 6 the book becomes slightly more technical, and is presumably addressed more directly to the specialist, although the clarity of style makes it quite suitable for the intelligent layman. A lengthy table details about 20 characteristics (sex, test scores, language difficulties, familial incidence) of 264 subjects, of whom 257 were deemed to be ' unambiguously dyslexic '. It is presented to ' set the scene ' for following chapters in which various aspects of the table are discussed in great detail, and for comparison with controls matched roughly for age and intelligence but with normal language facilities. Comparisons then show that the groups are indeed different in test profiles. Yet this seems to be an entirely unexceptional finding. Given that Miles had deemed these cases from his records to be ' unambiguously dyslexic ' he had already applied a set of criteria which distinguished them from ' normals '. In which case the tabulated data merely confirm that when he sets his mind to it he can find those he's looking for. Presumably there might well be a group of ' ambiguously dyslexic ' who might fill in all the statistical differences between his two groups, in which case the concept of dyslexia becomes a very idiosyncratic notion indeed.

Chapters 9 to 21 examine differences in a variety of test procedures in considerable detail. The author's experience highlights these with great clarity and they will be of interest to anyone working with language disadvantaged cases. It must, nevertheless, be remembered that the differences claimed by this study may very well be artefacts of the sampling procedure, and are in no way evidence for the existence of a distinct category of ' dyslexia '.

As a compendium of Professor Miles' wisdom and experience this book will be greatly appreciated by the layman and the educator. As a scientific contribution to the specialist it is very disappointing.

G. BROWN

TORY HIGGINS, E., RUBLE, D., and HARTUP, W. W. (Eds.) *Social Cognition and Social Development: A Sociocultural Perspective.* Cambridge: Cambridge University Press: no price.

This book contains a number of contributions stemming from a series of papers presented at a conference in 1979 and is primarily concerned with attempting to underline the *social* in the development of social cognition.

Following a general introduction by the editors the remaining 15 chapters are divided into three sections. The first of these deals with social understanding and inter-personal relations, the second with value internalisation and moral development while the four chapters in the final section provide a commentary and overview. It is this final section that emerges as one of the strengths of the book and it itself provides a much more detailed review of the more substantive content of the book than is possible here.

In his review chapter (the concluding one to the book but one perhaps best read immediately after the introduction) Trabasso sums up the main concern thus: ". . . part of the problem is that psychologists spend most of their time studying what goes on inside the head (cognition) and what children do to others (behaviour) and very little time on where children find themselves and what happens to them . . ." (p. 395). Likewise, Damen in his excellent review chapter suggests that many of the results reported in the book that relate directly to age differences " are difficult to interpret as we do not know enough about the social phenomena to which they apply " (p. 382). Two other review chapters by Turiel and by Maccoby make similar points.

The substantive chapters can be said to provide arguments for examining the social aspects of social development rather than actually giving examples of it. There are exceptions to this. Higgins and Parsons provide a good discussion of the ways in which changes in the structures and demands of (American) schools can be linked to age-related changes in the making of achievement and moral judgments. Ruble in a chapter on children's use of social comparison processes is keen to point to the possibly changing functions of social comparison with age and with the effects of changing demands made by others—for example, parents becoming increasingly concerned with relative rather than absolute standards of attainment.

In the section on moral development, Hoffman writes interestingly about the interaction between a child's memory for moral rules and the discipline techniques used by parents. The latter are held to influence the former which in turn mediate the latter. Lepper also provides a most interesting chapter on the relationship between social control processes and the internalisation of social values. His analysis makes extensive and elegant use of attribution theory and research, but, as Maccoby says in her review chapter which deals extensively with Lepper, such elegance may also tend to oversimplify the complex interrelationships within a family group.

On balance, the book promises rather more than it actually delivers. In my view it is clear that the study of the social side of social cognition requires extensive, detailed and thoughtful study of actual specific social contexts in order to demonstrate the relationship between them and a child's cognitions. Not all *changes* that take place in a child's thinking about social matters are necessarily examples of *development.* As a child moves from one social context to another different demands will be experienced. Some, but only some, of these changing demands will give rise to genuine development in that the child will come to acquire different ways of thinking—new cognitive processes. The psychologist's task is to identify those changing demands that have this property and then to go on to say why. After reading this book (which is still a recommended activity) the reader may well also be persuaded of this—but will still want to see it put into practice.

COLIN ROGERS

PUBLICATIONS RECEIVED

Books for review have been selected from the following

Alland, A. (1983). *Playing with Form.* New York: Columbia University Press, c. $38·50, p. $18·00.

Altman, I., and Wohlwill, J. F. (Eds.) (1983). *Behaviour and the Natural Environment.* Vol. 6 of ' Human Behaviour and Environment: Advances in Theory and Research'. New York: Plenum Press, $35·00.

Anderson, D., and Vervoorn, A. (1983). *Access to Privilege.* Canberra: Australian National Press, $15·50.

Anderson, L. W. (1983). *Time and School Learning.* Beckenham, Kent: Croom Helm, £17·95.

A.P.U. (1983). *Science in Schools Age 11.* Report No. 1 of Assessment of Performance Unit. London: Department of Education and Science.

Asso, D. (1983). *The Real Menstrual Cycle.* Chichester: Wiley, £6·95.

Atkinson, C. (1983). *Making Sense of Piaget: The Philosophical Roots.* Henley-on-Thames: Routledge and Kegan Paul, £11·95.

Australian Council For Educational Research (1983). *Philanthropic Trusts in Australia.* Victoria, Australia: ACER. (4th edit.)

Bain, B. (Ed.) (1983). *The Sociogenesis of Language and Human Conduct.* New York: Plenum Press, $50.

Bickman, L. (1983). *Applied Social Psychology Annual—4.* London: Sage, c. £21·75, p. £10·95.

Booth, T., and Potts, P. (1983). *Integrating Special Education.* Oxford: Blackwell, p. £5·95.

Bower, G. H., (Ed.) (1983). *The Psychology of Learning and Motivation.* Vol. 16. Advances in Research and Theory. New York, London: Academic Press, £21·60.

Bowles, G., and Klein, R. D., (Eds) (1983). *Theories of Women's Studies.* Henley-on-Thames: Routledge and Kegan Paul, £5·95.

Brent, A. (1983). *Philosophy and Educational Foundations.* London: Allen and Unwin, £7·95.

Cohen, D. (1983). *Piaget: Critique and Reassessment.* Beckenham, Kent: Croom Helm, c. £13·95, p. £6·95.

Cohen, L., and Manion, L. (1983). *A Guide to Teaching Practice.* London: Methuen, c. £12·50, p. £6·95.

Davey, A. (1983). *Learning to be Prejudiced—Growing Up in Multi-Ethnic Britain.* London: Edward Arnold, £8·95.

Delamont, S. (1983). *Interaction in the Classroom.* (2nd edit.) London: Methuen, £3·50.

Dixon, J. P. (1983). *The Spatial Child.* Illinois: Charles C. Thomas, $24·75.

Entwistle, N. (1983). *Understanding Student Learning.* Beckenham, Kent: Croom Helm, £14·95.

Everard, W., Hindley, C. B., Bot, A., and Van Der Werff Ten Bosch (Eds.,) (1983). *Development in Adolescence.* The Hague: Martinus Mojhoff, $43·50.

Fine, J., and Freedle, R. O. (Eds.) (1983). *Developmental Issues in Discourse.* Vol. 10 in series ' Advances in Discourse Processes '. New Jersey: Ablex Publishing Corporation, £29·95.

Fogelman, K. (1983). *Growing up in Great Britain.* London: Macmillan, £25·00.

Fordham, A. M. (1983). *The Context of Teaching and Learning: Report on the First Phase of the IEA Classroom Environment Study.* Victoria, Australia: ACER, $8·00.

Fowler, W. (1983). *Potentials of Childhood.* Vols. I & II. Aldershot: Lexington Books, £29·50 per volume.

Gale, J., and Marsden, P. (1983). *Medical Diagnosis.* Oxford: Oxford University Press, £17·50.

George, C. (1983). *Apprendre par l'Action.* Paris: Presses Universitaires de France, 110FF.

Gipps, C., Steadman, S., Blackstone, T., and Stierer, B. (1983). *Testing Children.* London: Heinemann Educational, £11·40.

Glendinning, C. (1983). *Unshared Care: Parents and their Disabled Child.* Henley-on-Thames: Routledge and Kegan Paul, £6·95.

Groner, R., Groner, M., and Bischof, W. F. (1983). *Methods of Heuristics.* London: Lawrence Erlbaum Associates, £37·95.

Irvine, S. H., and Berry, J. W. (Eds.) (1983). *Human Assessment and Cultural Factors.* New York: Plenum Press., $85·00.

Kratochwill, T. R. (Ed.) (1983). *Advances in School Psychology Vol. III.* New Jersey and London: Lawrence Erlbaum Associates, £29·95.

Kunick, P. (1983). *Relating to Learning.* Hemel Hempstead: Allen and Unwin, c. £12·00, p. £4·50.

Lewis, M. (Ed.) (1983). *Origins of Intelligence in Infancy and Early Childhood.* New York: Plenum Press ,$29·50.

Lindsey, G. (1983). *Problems of Adolescence in the Secondary School.* Beckenham, Kent: Croom Helm, c. £15·95, p. £7·95.

Lock, A., and Fisher, E. (1983). *Language Development.* Beckenham, Kent: Croom Helm, £5·95.

Longacre, R. E. (1983). *The Grammar of Discourse.* New York and London: Plenum Press, $45·00.

Mayall, B., and Petrie, P. (1983). *Child-minding and Day Nurseries: What Kind of Care?* London: Heinemann Educational/London Institute of Education, £6·50.

Miller, A. (translated by Ward, R.) (1983). *The Drama of the Gifted Child and the Search for the True Self.* London: Faber and Faber, cased £12·50, p. £2·95.

Miller, A. (translated by Hannum, H. and H.) (1983). *For Your Own Good: Hidden Cruelty in Childrearing and the Roots of Violence.* London: Faber and Faber, £8·50.

Nitko, A. J. (1983). *Educational Tests and Measurement: An Introduction.* London: Harcourt Brace Jovanovich, £16·50.

O'Connell, A. N., and Russo, N. F. (Eds) (1983). *Models of Achievement.* New York: Columbia University Press, c. $32·50, p. $16·50.

Olson, D. R., and Bialystok, E. (1983). *Spatial Cognition: the Structure and Development of Mental Representations of Spatial Relations.* London: Lawrence Erlbaum Associates, £22·95.

Osterlind, S. J. (1983). *Test Item Basis.* (Series: ' Quantative Applications in the Social Sciences '.) Beverley Hills: Sage Publications, £3·95.

Paris, S. G., Olson, G. M., and Stevenson, H. W. (Eds.) (1983). *Learning and Motivation in the Classroom.* New Jersey and London: Lawrence Erlbaum Associates, £29·95.

Peters, A. M. (1983). *The Units of Language Acquisition.* Cambridge: Cambridge University Press, c. £16·50, p. £6·95.

Piattelli-Palmarini, M. (1983). *Language and Learning: the Debate between Jean Piaget and Noam Chomsky.* Henley-on-Thames: Routledge and Kegan Paul, £5·95.

Richards, C., and Holford, D. (1983). *The Teaching of Primary Science: Policy and Practice.* London: The Falmer Press, c. £11·95, p. £6·50.

Robson, C. (1983). *Experiment, Design and Statistics in Psychology.* London: Penguin Books, £2·95.

Rogers, D., and Sloboda, J. A. (Eds.) (1983). *The Acquisition of Symbolic Skills.* New York: Plenum Press, $79·50.

Rosenthal, R., and Rosnow, R. (1984). *Essentials of Behavioural Research Methods and Data Analysis.* Maidenhead: McGraw Hill, £23·95.

Skinner, A. (1983). *Disaffection from School Issues and Interagency Responses.* Leicester: National Youth Bureau, £3·95.

Smith, F. (1983). *Essays into Literacy.* London: Heinemann Educational, £7·50.

Topping, K. (1983). *Educational Systems for Disruptive Adolescents.* Beckenham, Kent: Croom Helm, c. £14·95, p. £7·95.

Wainer, H., and Messick, S. (Eds) (1983). *Principles of Modern Psychological Measurement: A Festschrift for Frederic M. Lord.* London: Lawrence Erlbaum Associates, £29·95.

Watkins, M. J. (1983). *Multiple Choice Questions: Psychiatry.* Sevenoaks: Hodder and Stoughton Education, £2·95.

White, M. A. (1983). *The Future of Electronic Learning.* New Jersey and London: Lawrence Erlbaum Associates, £14·95.

Wolfendale, S. (1983). *Parental Participation in Children's Development and Education.* New York: Gordon and Breach Science Publishers, $38·50.

Zigler, E. F., *et al.* (1983). *Children, Families and Government.* Cambridge: Cambridge University Press, £25·00.

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

British Journal of Psychology
British Journal of Medical Psychology
Journal of Occupational Psychology
British Journal of Social Psychology
British Journal of Clinical Psychology
British Journal of Developmental Psychology
British Journal of Mathematical and Statistical Psychology
The British Journal of Educational Psychology

The *British Journal of Psychology*, the *British Journal of Medical Psychology*, *Journal of Occupational Psychology*, the *British Journal of Developmental Psychology*, the *British Journal of Social Psychology* and the *British Journal of Clinical Psychology* appear in quarterly parts. The 1984 subscription prices, payable in advance, are £56·00 (US$110·00) for the *British Journal of Psychology*; £36·00 (US$64·50) for the *British Journal of Social Psychology*; £35·00 (US$64·50) for the *British Journal of Clinical Psychology*; £33·00 (US$64·50) for the *British Journal of Developmental Psychology*; £46·00 (US$90·50) for the *British Journal of Medical Psychology*; and £39·00 (US$79·50) for the *Journal of Occupational Psychology*.

The *British Journal of Mathematical and Statistical Psychology* is issued twice a year, in May and November. The subscription rate, payable in advance, is £39·00 per volume post free (US$79·50 in the USA).

Subscriptions to any of these journals may be sent to a bookseller or subscription agent, or to The British Psychological Society, The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK. A special price of £257·00 (US$500·00) for all seven journals going to the same address is offered for 1984.

The British Journal of Educational Psychology is issued three times a year, in February, June and November. The prepaid subscription is £15·00 a year, post free, or £5·50 each number. The prepaid subscription price in the USA is $40·00 post free. Subscriptions should be sent to the Scottish Academic Press, 33 Montgomery Street, Edinburgh EH7 5JX.

MANUSCRIPTS

For the *British Journal of Psychology* should be sent to Professor D. E. Blackman, Department of Psychology, University College Cardiff, PO Box 78, Cardiff CF1 1XL.
For the *British Journal of Medical Psychology* should be sent to Dr Sidney Crown, Department of Psychiatry, The London Hospital, Whitechapel, E1 1BB.
For the *Journal of Occupational Psychology* should be sent to Professor P. Warr, MRC/SSRC Social and Applied Psychology Unit, Department of Psychology, The University, Sheffield S10 2TN.
For the *British Journal of Social Psychology* should be sent to Dr A. S. R. Manstead, Department of Psychology, The University, Manchester M13 9PL.
For the *British Journal of Clinical Psychology* should be sent to Dr D. Shapiro, MRC/SSRC Social and Applied Psychology Unit, Department of Psychology, The University, Sheffield S10 2TN.
For the *British Journal of Mathematical and Statistical Psychology* should be sent to Dr C. Leach, Department of Psychology, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU.
For the *British Journal of Developmental Psychology* should be sent to Professor P. E. Bryant, Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD.
For *The British Journal of Educational Psychology* should be sent in duplicate to Professor Geoffrey Brown, School of Education, University of East Anglia, Norwich NR4 7TJ.

Contributors are advised to consult *Suggestions to Contributors* issued by the Standing Committee on Publications of the British Psychological Society and obtainable, price £2·00 post free, from The British Psychological Society, St. Andrew's House, 48 Princess Road East, Leicester LE1 7DR.

All manuscripts should be submitted in duplicate.

OFFPRINTS

Contributors on application receive fifty copies of their papers free. Additional copies may be obtained in units of fifty copies.

REPRINTS

The *Journal* is available in reprint from Swets and Zeitlinger B.V., P.O. Box 810, 2160 SZ Lisse, The Netherlands; Vols. 21-29, 31-47.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £30·00.

They are entitled to purchase any or all of the Society's Journals at the special rates of £6·00 per Journal per annum.

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, St. Andrew's House, 48 Princess Road East, Leicester LE1 7DR.

BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY,
VOL. 54, PART 2, 1984

CONTENTS

ARTICLES

HACKER, R. G. — A Hierarchy of Intellectual Development in Science . 137

IVES, S. W. — The Development of Expressivity in Drawing . . 152

CUNNINGHAM, C. C. and SLOPER, P. — The Relationship between Maternal Ratings of First Word Vocabulary and Reynell Language Scores . . . 160

PHILLIPS, C. J. — Use of Abbreviated Cognitive Tests in the Assessment of Children's Special Educational Needs . . . 168

CHAPMAN, J. W. and LAWES, MERYL M. — Consistency of Causal Attributions for Expected and Actual Examination Outcome: A Study of the Expectancy Confirmation and Egotism Models 177

SJÖBERG, L. — Interests, Effort, Achievement and Vocational Preference. 189

MARSH, H. W. — Experimental Manipulations of University Student Motivation and their Effects in Examination Performance . 206

HARGIE, O. D. W. — Training Teachers in Counselling Skills: The Effects of Microcounselling 214

RESEARCH NOTES:

NASH, B. CHRIS. — Kindergarten Task Management and Grade 2 Perceptions of Teacher-Learner Roles 221

ERWIN, P. G. and CALEV, A. — The Influence of Christian Name Stereotypes on the Marking of Children's Essays 223

O'NEIL, M. J. and CHILD, D. — Biggs' SPQ: A British Study of its Internal Structure . 228

MCINTYRE, T. C. — The Relationship Between Locus of Control and Teacher Burnout 235

BOOK REVIEWS 239

PUBLICATIONS RECEIVED 243

23 NOV 1984

UK ISSN 0007-0998

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

NOVEMBER 1984

VOL. 54, PART 3

THE
BRITISH JOURNAL
OF
EDUCATIONAL PSYCHOLOGY

Edited by
Professor GEOFFREY BROWN

Assistant Editors:
Dr. R. L. BURDEN
Dr. C. W. DESFORGES
Dr. D. FONTANA

Sub-Editor:
Dr. S. CUNNINGHAM

Book Review Editor:
Dr. C. W. DESFORGES

Advised by the Editorial Committee:
Dr. H. BELOFF
Professor M. B. SUTHERLAND
Professor J. WRIGLEY

Vol. 54, Part 3

The British Journal of Educational Psychology is issued by the British Psychological Society once a term, in February, June and November. The subscription is £15·00 a year post free, payable in advance, or £5·50 each number.

Subscriptions and Orders should be sent to Scottish Academic Press, 33 Montgomery Street, Edinburgh, EH7 5JX. Members of the British Psychological Society receive the *Journal* on special terms.

Papers for publication in duplicate should be sent to Professor Geoffrey Brown, School of Education, University of East Anglia, Norwich NR4 7TJ. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for *The British Journal of Educational Psychology* should be sent to: Dr. C. W. Desforges, School of Education, University of East Anglia, Norwich NR4 7TJ.

Advertisements for this *Journal* should be sent to Scottish Academic Press.

Published by
SCOTTISH ACADEMIC PRESS, 33 MONTGOMERY STREET, EDINBURGH, EH7 5JX,
for the British Journal of Educational Psychology, Ltd.

All rights reserved

© 1984. The British Journal of Educational Psychology

PRINTED IN GREAT BRITAIN BY CLARK CONSTABLE, EDINBURGH, LONDON, MELBOURNE

Br. J. educ. Psychol., 54, 245-253, 1984

NUMBER UNDERSTANDING IN PRESCHOOL CHILDREN

BY SUSAN MICHIE
(*Department of Experimental Psychology, University of Oxford*)

SUMMARY. Judgments of relative number made by 4-year-old children were assessed by asking them to count and compare two collections of counters under various conditions. Counters were presented either in heaps, to be counted one at a time into opaque boxes, or to be placed on top of dots in rows. In the experimental series a visual display of the number sequence was provided, and the appropriate number marked after each collection had been counted. To control for the length cues in the visual display, a further test was carried out using a visual display with no numbers on it. A condition with no visual display was also used. Performance was significantly more accurate when numbers and numerical dot patterns were visible and when the counters were counted into boxes rather than being arranged into rows. This suggests that children's poor performance on relative number judgment tasks may be due not to a poor understanding of number but rather to memory factors and to over-reliance on perceptual cues. Aiding their memory of the number sequence and the numbers counted, and removing perceptual cues, enabled them to use their counting skills effectively.

INTRODUCTION

PIAGET (1952) concludes that children who fail his conservation test are limited to judging numerosity in terms of perceptual cues such as length and density. This conclusion has been generally accepted, despite its never having been tested directly (e.g., by Pufall and Shaw, 1972; Smither *et al.*, 1974; Baron *et al.*, 1975; Brainerd, 1977). When numbers are small, children have been shown to focus spontaneously on the numerosity of a set (Gelman, 1972). There is also evidence that 4- to 6-year-olds use counting to compare number before they perform correctly on the number conservation task (Saxe, 1979).

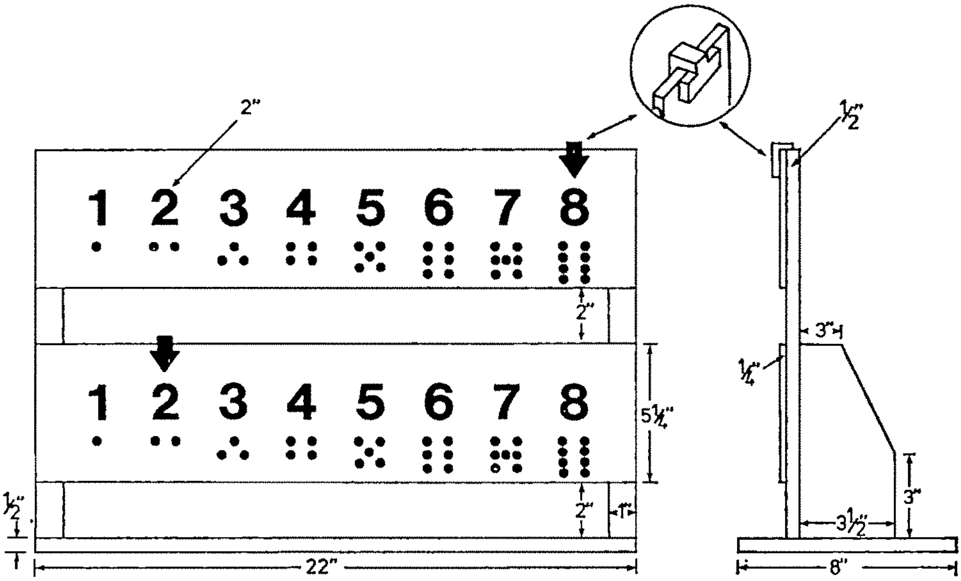
This study does not seek to establish whether children have a ' true ' understanding of number. Rather, it looks at the question of whether young children gain numerical information from their counting behaviour and use that information appropriately in making judgments. Gelman and Gallistel (1978) have shown that children use certain counting principles even when their counting is inaccurate. The experiment reported here does not investigate *how* children count. Instead, it looks at the significance that children attach to their counting.

The traditional task used to assess children's number judgments involves presenting two rows of items that may differ in length or density and asking for judgments of relative numerosity. This procedure presents two problems for present purposes. The first is that the child is presented with perceptual cues which frequently may be misleading (the lengths and densities of the rows) at the same time as he/she is asked for judgments of their relative number. Given the opportunity to use a length cue, the child will take it. This tells one nothing about the child's use of number in the absence of a length cue. The second problem is that *using* counting information requires remembering counting information. In order to make a relative number judgment, children must remember the final numerals for each row and where each of these numerals lies in the number sequence.

The present experiment aims to assess children's number understanding directly, by making two innovations to the number judgment task. One is the exclusion of perceptual cues (counters are counted into opaque boxes). The second is the inclusion of a visual display of the number sequence to aid the children's memory (Figure 1). As the children count a collection of items, the arrow is moved to the numeral corresponding to the one spoken by the child. When the number judgment is asked for, the child has in front of him/her a visual representation of the number (by symbol and dot pattern) of both collections.

FIGURE 1

THE VISUAL DISPLAY OF THE NUMBER SEQUENCE



Thus, in the traditional task, children who understand number may be failing because they are given the opportunity to use the length cue or to forget the number counted. The present experiment removes these possibilities: failing the task in this experiment would, indeed, suggest a lack of understanding about number. It is unlikely that children not able to use symbol, oral or visual cues for number in this task would show an understanding of number in other tasks.

If a preference for a length cue and difficulty in remembering the number sequence do account for the poor performance in the traditional number task the following results would be expected in this experiment. One would expect children to be most accurate when counting into opaque boxes and when the visual display of the number sequence is present. One would expect children to be least accurate when counting rows (with their distracting length and density cues) and when no visual display is present. The presence of one or the other of these conditions should lead to an intermediate level of accuracy.

The present experiment serves two purposes. Firstly, it aims to establish whether children do demonstrate an understanding of number when conditions are optimal. Secondly, it aims to determine whether the poor performance in the traditional number task is due to a preference for the length cue and/or difficulties in remembering.

METHOD

Subjects

Forty-eight 4-year-old children who were able to count to eight formed the sample. Twenty-six were from one nursery school in the Oxford area, 22 from another. There were 24 girls and 24 boys. The mean age was 4 years 6 months (range 4 years 0 months to 4 years 11 months). The group was subdivided into the 24 youngest and the 24 oldest. The mean age of the young group was 4 years 4 months (range 4 years 0 months to 4 years 7 months) and the mean age of the old group was 4 years 9 months (range 4 years 7 months to 4 years 11 months).

Design

Forty-eight children were given a number judgment task on a total of 24 trials: four trials in each of two conditions under each of three procedures. In one condition, children counted while placing counters on two rows of each of the cards (Figure 2); in the other condition, children counted equivalent numbers of counters from piles into two boxes. Each condition was presented with three different visual aid procedures on different occasions:

(1) Counting with the aid of a visual display with numbers (experimental);
(2) Counting with the aid of a visual display without numbers or dot patterns (control 1);
(3) Counting with no visual display.

In procedure (2), the arrows were moved in similar manner to procedure (1). This served as a control to ensure that any advantage provided by the visual display in procedure (1) accrued from number information and not simply from the distance moved by the arrow.

The four trials in each condition included two in which equal collections were compared and two in which unequal collections were compared. In the case of the cards with equal collections, one trial had rows of equal length and density and the other had rows of different length and density. For the cards with unequal collections, one trial had rows of either the same length or the same density and the other had rows of different length and density.

The design was repeated-measures, all the children doing all six tasks. The sequence in which the six tasks were presented to each child was counterbalanced. The order of the three visual display procedures was arranged by a Latin Square design. Within each of these, half the children were given the cards before the boxes and half were given the boxes before the cards. Within each condition, the order in which the four counting trials were given was randomised, and different for each child.

Materials

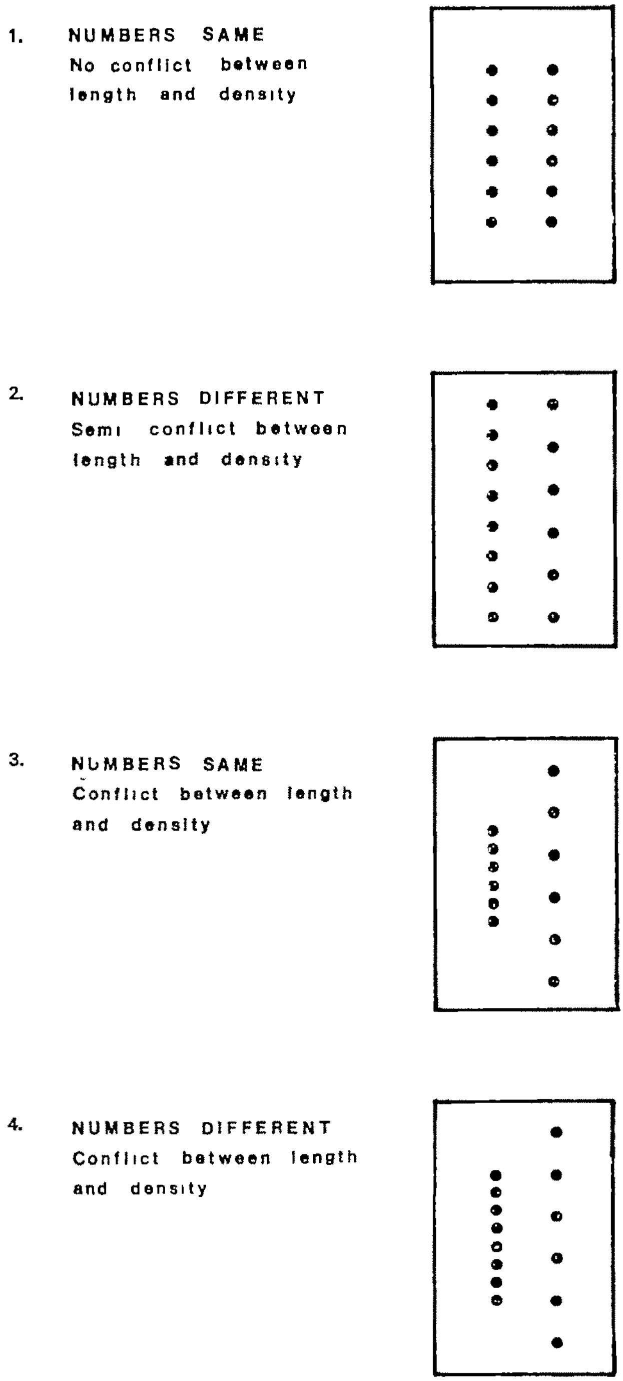
Four white cards, 30 cm high and 20 cm wide, showing two rows of pencil dots (Figure 2) were used. Two cards had rows with the same numbers and two with different numbers (six and eight). Of each of these pairs, one card had rows with two cues to number, length and density, conflicting (one row was longer and the other more dense). On each of the pencil dots, children placed red or green plastic counters, 0·6 cm in diameter.

The two opaque plastic tubs into which subjects put the counters under the condition of no perceptual cues were 10 cm in diameter and 7 cm high, one painted red and the other green. Each had a slit cut out of its lid.

A visual display of the number sequence is shown in Figure 1. Constructed out of white perspex, it has green numerals and dot patterns on the upper scale and red on the lower. The movable perspex arrows were correspondingly green and red. The display was 56 cm wide and 38 cm high.

FIGURE 2

THE ARRAYS USED


1. NUMBERS SAME
No conflict between length and density
2. NUMBERS DIFFERENT
Semi conflict between length and density
3. NUMBERS SAME
Conflict between length and density
4. NUMBERS DIFFERENT
Conflict between length and density

Procedure

Children were shown two small dolls, ' Mr. Red ' and ' Mr. Green ', and were told: " I'm going to give Mr. Red some buttons and Mr. Green some buttons and I want you to tell me whether you think they have the same number or whether one has more."

After three consecutive correct trials, they were given the following instructions:

Cards condition

" Now I'm going to show you some cards with two rows marked. I want you to put the green buttons onto the pencil dots of Mr. Green's row and the red buttons onto the pencil dots of Mr. Red's row. As you do that I want you to count how many each of them have so that you can tell me whether you think they have the same number or whether one has more."

After children had counted the second row of counters, they were told: " Good, Mr. Red has six (or eight) buttons and Mr. Green has six (or eight) buttons. Do you think they have the same number or do you think one has more? "

The order of counting was varied randomly between left and right, red and green, six and eight.

Boxes condition

" Now I'm going to give a green box to Mr. Green and a red box to Mr. Red. I want you to put the green buttons into Mr. Green's box, counting them as you do so and then put the red buttons into Mr. Red's box, counting them as you go. Afterwards I want you to tell me whether you think they have the same number or whether one has more." The rest of the trial was the same as with the cards.

These Procedure (3) instructions were preceded by the following for:

Procedure (1)*: the visual display with numbers.*

The children were asked if they recognised what was written on the visual display. If they did, they were asked to recite the numbers. If they were unsure, they were told the numbers associated with the numerals and dot patterns. They were then told:

" As you count the green buttons, I'm going to move the green arrow along to the same number that you count so you can see what number you've counted to. When you count the red buttons I'll move the red arrow along to the same number that you count."

The arrow was moved discontinuously to correspond with the children's counting. When the final question was asked, the experimenter pointed to the position of the green arrow when its number was repeated by the experimenter and pointed to the position of the red arrow when its number was repeated.

Procedure (2)*: the visual display without numbers.*

" This time I'm going to time how long you take to count. When you count the green buttons I'll slide the green arrow along and stop when you stop. We can then see how long you take to count both lots of buttons."

The explanation that the arrow was connected with timing was given to provide some rationale for sliding the arrow. The experimenter slid the arrow along continuously at a suitable rate so that, however long the child took, the arrow ended up at the ' six ' position (although the numbers were covered) when six was counted and at the ' eight ' position when eight was counted.

Before the final question was asked, the experimenter said: " Mr. Red has six (or eight) and Mr. Green has six (or eight) and you took this long to count the red buttons and this long to count the green buttons " and pointed to the respective red and green arrows.

Children were seen on three separate occasions, separated by one to two days. The number and type of errors were recorded. Children were seen on their own in the staff room after the experimenter felt she had become sufficiently acquainted with them. All the children appeared to enjoy the ' game '.

RESULTS

The mean number of errors in each of the conditions is shown in Table 1. The mean number of errors is smaller when children counted into boxes than on cards. It is also smaller when children were shown numbers on the visual display. The presence of numerals increased the accuracy over both controls (i.e., when the visual display was absent and when it was present but only indicating the distance moved by the arrows).

TABLE 1

MEAN ERRORS (OUT OF A POSSIBLE 2 PER CELL). STANDARD DEVIATIONS ARE IN BRACKETS.

		Condition			
		Cards		Boxes	
Procedure	Age	Same Number	Different No.	Same Number	Different No.
Visual display with nos.	Young	0·42 (0·58)	0·62 (0·77)	0·12 (0·44)	0·12 (0·44)
	Old	0·37 (0·57)	0·29 (0 62)	0·17 (0·56)	0·00 (0·00)
Visual display without nos.	Young	0·67 (0·64)	0·87 (0·79)	0·33 (0·70)	0·33 (0·56)
	Old	0·50 (0·78)	0·79 (0·88)	0·21 (0·59)	0·17 (0·38)
No visual display	Young	0·71 (0·69)	1·17 (0·81)	0·50 (0·78)	0·50 (0·72)
	Old	0·58 (0·71)	1·00 (0·93)	0·37 (0·77)	0·46 (0·66)

Unequal collections are more difficult than equal collections in the case of cards, but there does not appear to be a difference in the case of boxes. As the task becomes more difficult (with less number information), this difference increases. In 11 out of 12 cells, the young 4-year-olds made more errors than the old ones.

In an ANOVA, children were divided into younger and older age groups, giving one between-subject variable. There were three within-subject variables: ' visual display procedure ', with three levels (' with numbers ', ' without numbers ', and ' absent '), ' card *versus* box condition ', with two levels; and ' same number *versus* different number trials ' with two levels.

There was a significant difference between the three visual display procedures ($F(2,92) = 23{\cdot}63$, $P<0{\cdot}001$) and between the box and card conditions ($F(1,46) = 55{\cdot}30$, $P<0{\cdot}001$). A Tukey post-hoc test showed that children performed significantly better when the visual display was present, with numbers than without ($P<0{\cdot}001$) and better when the display indicating distance only was present than when it was absent ($P<0{\cdot}001$). The ANOVA also confirmed the prediction that children would be significantly more accurate in the ' box ' than in the ' card ' condition ($F(1,46) = 22{\cdot}17$, $P<0{\cdot}001$). There was no significant difference between age groups.

There was one significant interaction: between ' box *versus* card ' and ' same *versus* different numbers ' ($F(1,46) = 9{\cdot}05$, $P<0{\cdot}01$). Simple main effects analysis showed that there was an insignificant difference between judgments of same and different collections when counted into boxes ($F = 0{\cdot}10$) but a highly significant one when counted onto cards ($F = 20{\cdot}41$, $P<0{\cdot}001$), with greater accuracy on the ' same number ' than ' different number ' cards. This may be due to the nature of the displays on the cards. Both the ' different number ' displays contain conflict between relative length and number, whereas only one of the two ' same number ' displays does. If children are using length to judge number, one would therefore expect greater accuracy in these ' same number ' than ' different number ' displays.

The increased accuracy of counting into boxes rather than on cards was significant both when numbers were the same ($F = 84{\cdot}80$, $P < 0{\cdot}001$) and when different ($F = 25{\cdot}03$, $P < 0{\cdot}001$).

There is a large difference in errors for the traditional task (cards and no visual display) compared with the two innovations combined (boxes and visual display). The total errors show this clearly. Of a possible 96 errors, 82 were made in the traditional task and only 10 in the new task. The new task included two innovations—a display of the number sequence and the removal of misleading length and density cues. When only one innovation was included, the number of errors was intermediate: 41 when the visual display was presented (without the boxes) and 44 when boxes were used (without the visual display). These results demonstrate that the nature of the traditional task discourages children from using number in making their judgments. The author's task facilitates numerical judgments based on counting.

Performance on cards was analysed by comparing conflict and non-conflict arrays rather than comparing size of collection counted. Mean errors are shown in Table 2. The analysis was by an ANOVA with one between-subject variable, age, and two within-subject variables, visual display procedure (three levels) and conflict *versus* non-conflict (two levels).

TABLE 2

MEAN ERRORS (OUT OF A POSSIBLE 2 PER CELL) WITH CARD ARRAYS DIVIDED INTO NON-CONFLICT AND CONFLICT (STANDARD DEVIATIONS ARE IN BRACKETS)

		Cards	
Procedure		Non-Conflict	Conflict
Visual display with numbers	Young	0·33 (0·48)	0·75 (0·74)
	Old	0·12 (0·34)	0·54 (0·83)
Visual display without numbers	Young	0·37 (0·57)	1·16 (0·76)
	Old	0·45 (0·58)	0·71 (0·80)
No visual display	Young	0·58 (0·58)	1·37 (0·82)
	Old	0·54 (0·66)	1·00 (0·88)

There was a significantly higher error rate in the conflict than non-conflict arrays ($F(1,46) = 51{\cdot}57$, $P < 0{\cdot}001$). Tukey post-hoc tests showed this to be the case in all procedures, even when children were shown the number sequence. This shows that perceptual cues of the rows were influencing children's judgments even when number information is readily available.

DISCUSSION

The number of errors of judgment after counting was significantly reduced when the visual display showed numbers than when it did not, showing that children must be using number information ' per se ' to aid their judgment. This number information may not derive from the specific numerals; there is the possibility that it may be derived from the number of times that the arrow stops at a numeral. The significant difference between the two controls shows that children are making use of the relative distances the arrow moved, as would be expected.

Children are also significantly more accurate when they are putting counters into boxes than when they are placing them in rows on cards. They rely more on information from counting when alternative perceptual cues are not present. When perceptual cues are present they are preferred, probably because they are visible and do not have to be remembered before they can be compared. Either eliminating perceptual cues or aiding memory for number counted reduces errors significantly. If both are combined, performance is virtually error-free. Number counted can form the basis for these 4-year-olds' number judgments.

Piaget's assertion that young children lack a concept of number if they invariably judge arrays by length is questioned by this experiment. Under appropriate conditions children can be shown to conceive of number as a dimension separable from dimensions of length and density.

The belief arising from Piaget's conclusions, that counting by young children is a rote and meaningless task (e.g., Isaacs, 1960; Ginsburg, 1977), appears to be wrong. Although children may prefer in some situations to base their judgments on information other than number, nevertheless they may understand the significance of the numerals that they recite. The results of their counting may be played down by children, not because they lack an understanding of number, but because they judge it to be less reliable than other sources of information relevant to number (e.g., length of row). Children have been found to count more frequently after feedback showing them that numeral counted is a more reliable cue to number than is length of row (Michie, 1982). If this is the explanation, teachers should be more concerned with teaching children how to assess in which situations which strategies (including counting) are likely to be most reliable. Teaching the conceptual nature of numbers recited may be redundant.

The explanation that children use length because they think that counting is unreliable raises other questions. Although they understand the significance of number as counted in some situations, the reluctance to use it in others suggests understanding to be partial and situation-dependent. More research is needed on how variations in children's readiness and ability to use counting change with variations in the situation.

It is clear that children are making use of some kind of information that they derive from the number system. This raises the question of what aspect or aspects of number it is that children are using. Bryant (1974) argues that number has both a relative (" which has more? ") and an absolute (" how many? ") component. The relative aspect refers to the position of the number in the sequence and the absolute aspect refers to the actual amount of the number. Absolute number appears to develop before the relative concept (Michie, 1984).

ACKNOWLEDGMENTS.—This study was carried out as part of a doctoral thesis and I thank Peter Bryant for supervising the work, John McShane for help in writing this paper, and a Pirie Reid scholarship for financial support.

Requests for reprints should be sent to Dr. Susan Michie, Psychology Department, 2nd floor, The Royal Free Hospital, Pond Street, London NW3 2QG.

REFERENCES

BARON, J., LAWSON, G., and SIEGEL, L. S. (1975). Effects of training and set size on children's judgments of number and length. *Dev. Psychol.*, 11, 583-588.

BRAINERD, C. J. (1977). Effects of spatial cues on children's cardinal number judgments. *Dev. Psychol.*, 13, 425-430.

BRYANT, P. E. (1974). *Perception and Understanding in Young Children.* London: Methuen.

GELMAN, R. (1972). Logical capacity of very young children: number invariance rules. *Child Dev.*, 43, 75-90.

GELMAN, R., and GALLISTEL, C. R. (1978). *The Child's Understanding of Number.* Cambridge, Mass.: Harvard University Press.

GINSBURG, H. (1977). *Children's Arithmetic.* New York: D. Van Nostrand.

ISAACS, N. (1960). *New Light on Children's Idea of Number: The Work of Professor Piaget.* London: Ward Lock Educational.

MICHIE, S. (1982). Cognitive uncertainty and number development in young children. D.Phil. Thesis, University of Oxford.

MICHIE, S. (1984). The development of absolute and relative concepts of number in preschool children. *Dev. Psychol.*, in press.

PIAGET, J. (1952). *The Child's Conception of Number.* London: Routledge and Kegan Paul.

PUFALL, P., and SHAW, R. (1972). Precocious thoughts on number: the long and the short of it. *Dev. Psychol.*, 7, 62-69.

SAXE, G. B. (1979). Developmental relations between notational counting and number conservation. *Child Dev.*, 50, 180-187.

SMITHER, S. J., SMILEY, S. S., and REES, R. (1974). The use of perceptual cues for number judgment by young children. *Child Dev.*, 45, 693-699.

(*Manuscript received* 24*th February*, 1984)

CALL FOR NOMINATIONS

NCME Award for the Best Technical Contribution to Educational Measurement

The National Council on Measurement in Education is intending to make an award this year (1984-85) for an outstanding technical *or* scientific *contribution to the field of educational measurement.*

Examples of technical contributions include, but are not limited to, innovative ways of solving practical or theoretical measurement problems, inventive instrument development techniques, creative testing procedures or products, and scientific contributions to measurement research methodology. Selection criteria are quality, innovativeness, and positive impact of the technology on the field of educational measurement.

To be eligible for this award, the technical contribution must have occurred initially during 1982, 1983, or 1984. One may either nominate his/her technical contribution or, with permission, someone else's. Those responsible for the technical contribution need not be members of NCME. To be considered for the award, four copies of a 3-5 page statement describing the technology, application area, and products or results of the effort should be sent to: Awards Committee Chair, c/o Anthony J. Nitko, Department of Educational Research Methodology, University of Pittsburgh, Pittsburgh, PA 15260. Later, finalists may be requested to submit additional information. The deadline for submission is February 1, 1985.

The award will be presented at the NCME's 1985 annual meeting in Chicago.

Br. J. educ. Psychol., 54, 254-264, 1984

THE MATHEMATICAL ACHIEVEMENTS OF DEAF CHILDREN FROM DIFFERENT EDUCATIONAL ENVIRONMENTS

By HEATHER A. WOOD, DAVID J. WOOD, MARIAN C. KINGSMILL, JACK R. W. FRENCH and S. PATRICIA HOWARTH
(*Deafness Research Group, Psychology Department, Nottingham University*)

Summary. The performances of 465 hearing and 540 hearing-impaired school-leavers on the Vernon-Miller Graded Arithmetic-Mathematics test are reported. A previous study revealed that hearing-impaired children in special schools and partially hearing units scored approximately three years below their hearing peers. Somewhat paradoxically, however, degree of hearing loss was not a major determinant of performance in the hearing-impaired sample. One possible explanation for these findings is that children in special education follow a different and less demanding curriculum than hearing children. To explore this possibility, 135 hearing-impaired children attending ordinary schools were tested. Multiple regression analysis reveals that while this ' mainstreamed ' sample performed better than those in special education, educational placement *per se* is not an important determinant of achievement. No sex differences emerge for any hearing-impaired group and hearing loss remains a significant ($r = -0{\cdot}26$, $P < 0{\cdot}001$) but weak predictor of performance.

INTRODUCTION

In an earlier study (Wood *et al.*, 1983a) the performances of 15-16 year-old hearing, deaf and partially hearing children on the Vernon-Miller Graded Arithmetic-Mathematics Test (1976) were compared. The children involved were tested in 1981 during their final year at school. An attempt was made to test every deaf and partially hearing school-leaver in special schools and partially hearing units in England and Wales, and 414 hearing-impaired children were involved. In addition, a sample of 465 hearing children drawn from three comprehensive schools was tested to provide both an up-to-date comparison group and detailed information about scoring patterns and errors for future analysis.

Two main results of our previous study are of interest here. First, it was found that degree of hearing loss (and, by implication, linguistic ability) was not highly correlated with mathematical attainment in the deaf and partially hearing, accounting for only 1·7 per cent of the variance in achievement (significant at the $P < 0{\cdot}008$ level). Related to this result was the finding that the performance of 271 deaf children in special schools (mean hearing loss 92 dB) was not significantly different from that of 143 children in partially hearing units (mean hearing loss 68 dB) even though the difference in hearing loss was highly significant ($P < 0{\cdot}00000$). A second finding was that both hearing-impaired samples scored significantly lower than the hearing controls (mean maths ages of 12·1, 12·8, and 15·5 years respectively). These results are somewhat paradoxical. If degree of hearing loss is not a major predictor of mathematical attainment, it is difficult to explain why deafness *per se* should lead to such a significant decrement in performance in comparison with hearing children.

One possible explanation for this discrepancy, which is explored in this paper, is that the difference between hearing-impaired and hearing children results not from deafness *per se* but from educational experience. If, for example, deaf and partially hearing children receive less instruction in mathematics than hearing children, or receive different types and quality of instruction from their teachers, then it would follow that their performance might differ from those of hearing children. One way to test this hypothesis is to examine the performance of hearing-impaired children in ordinary schools—children who are being taught mathematics alongside hearing peers. If these ' mainstreamed ' children perform similarly to hearing children and significantly better than children from special schools and units (with the effects of hearing loss partialled out) then the hypothesis that differential educational experiences explain deaf-hearing differences would gain credibility. We would add, however, that such evidence would not provide unequivocal evidence for this hypothesis for a number of possible reasons. We have no information about the selection policies and practices of different LEAs. Thus, we cannot assume that hearing loss is the only factor that influences school placement. Other important correlates of mathematical attainment such as intelligence, aetiology of deafness and home background (see Jensema, 1975) might well play a determining role in school placement and hence confound the relationships between placement, hearing loss, and attainment. Whilst not definitive, however, evidence in support of the hypothesis would inspire some confidence that further, more detailed and extensive studies of the performance of children in different educational settings might prove of value in evaluating the effectiveness of different placement policies with regard to hearing-impaired children and their mathematical achievements.

A negative result would be equally important. If school environment is *not* an important correlate of mathematical attainment, then it would suggest that other factors associated with deafness explain the relatively low attainment of hearing-impaired children. One such explanation has already been formulated and discussed (Wood *et al.*, 1983a; Wood *et al.*, 1983b).

Although we have argued against the hypothesis that spoken language ability is directly involved in performance on the test used on the grounds that hearing loss is highly correlated with measures of language performance (e.g., Conrad, 1979) but not with mathematics (see also Jensema, 1975, and Suppes, 1974), it is still possible that an indirect link exists between language, communication and the learning of mathematics—what we have termed the ' Indirect Linguistic Hypothesis ' (Wood *et al.*, 1983b).

Our argument proceeds as follows: if language and subvocalisation provide the major symbol system for mathematical operations, then we would expect high correlations between hearing loss and mathematical achievement. To date, evidence exists against such an hypothesis. However, it is possible that the actual teaching of mathematics and the transmission of concepts and procedures *is* related to receptive linguistic ability. Perhaps the pace of lessons and the child's ease of learning is constrained by the quality of teacher-child communication, leading to slower or more difficult learning of mathematics when a child has hearing problems. Thus, whilst a child might not think in words while doing mathematical problems, it is possible that linguistic problems still exert an indirect influence on his learning. Such an hypothesis is speculative and, in some ways, weak, for it would still seem to predict high correlations between children's receptive language ability (correlated with hearing loss) and mathematical competence.

The main aim of the current investigation of mathematical achievement of hearing-impaired children in ordinary schools is to shed new light onto these issues.

A second aim of the present investigation is to provide information about the audiological characteristics and mathematical abilities of hearing-impaired children in ordinary schools. Although there is considerable interest in the issue of school placement for handicapped children and an increasing desire, following the Warnock report and the 1981 Education Act, to see more handicapped children in ordinary schools, there is little information about the current achievements of hearing-impaired children in normal schools. Thus, any changes in placement policy and their effects are almost certain to proceed without systematic monitoring or evaluation. This investigation provides some information about the numbers, aetiologies and attainments of such children in England and Wales who left school in the summer of 1982. Although we cannot claim that our sample is complete, we do have available a list of the LEAs who participated in this study and various details of the children studied. In future, such information could be used, for example, to chart any changes in the composition of this population—to see if more severely hearing-impaired children are being educated in ordinary schools and to compare their performances in mathematics with our current samples. Such information is incomplete and less than ideal, but likely to prove of some value given the current dearth of relevant evidence.

METHOD

Subjects

Detailed information about the 414 deaf and partially hearing children who participated in the earlier study, whose performances will be compared with the mainstreamed children in this paper, have already been documented (Wood *et al.*, 1983a), as have details of the hearing group. One minor change in this sample should be noted, however. Nine children with mean hearing losses less than 30 dB in their better ear were included in the original study. Since the cut-off for inclusion in the current sample is 30 dB, these are excluded from the analyses. This leaves a total sample of 405 children in special education.

In an attempt to recruit every hearing-impaired school-leaver in ordinary schools, every LEA in England and Wales was contacted by mail and asked to provide information about the sex, hearing loss, aetiology, and school address of every eligible child in their area. To be included, a child must have a hearing loss of at least 30 dB averaged over five frequencies in the better ear. Sixty-four (69 per cent) of LEAs responded. On following up the schools named 183 (79 per cent) responded, to yield a total of 135 eligible children (81 girls and 54 boys). In the absence of any register of relevant children in England and Wales, it is not possible to say what proportion of the total population this sample represents nor to claim that it is necessarily representative of that population.

Mathematics test

The main criterion governing the selection of the test was the absence of extensive written language in its questions so as to minimise the demands for reading ability which might confound our measures of mathematical performance. We also needed a test applicable to a wide range of ability and one unlikely to have been seen by children before. The test used was the Graded Arithmetic-Mathematics Test developed by Vernon and Miller (1976). The test offers two forms—Junior and Senior—and has a time limit of 30 minutes.

Procedure

During the spring and summer terms of 1982 each school which had been identified by LEAs was sent copies of both Senior and Junior versions of the test and asked to give the form felt to be most suitable to the child's likely level of ability. Teachers were asked to administer the test following normal test instructions, using the optional pretest and ' oral ' presentation if they considered them applicable. This procedure duplicated that used in the initial study with both hearing-impaired and hearing children. As we already had information on the children's age, sex, and hearing loss, teachers were merely asked to check that the information was as up-to-date as possible.

Scoring

The total number of correct answers on the test gives each child a raw score which is then converted into a maths age on the basis of tables provided with the test. However, these tables do not cover all possible raw scores and have a ceiling of a maths age of 17·1 years. Problems of extrapolation thus arise when we try to arrive at an average maths age for the various samples. Averaging procedures based on both raw scores and derived maths ages are thus given for each sample. All other analyses, such as correlations between maths ages and hearing loss, are calculated from children's maths age scores.

RESULTS

Hearing status and mathematical ability

As reported in the previous study, the average maths age for the sample of 465 hearing children was 15·5 years (15·6 using raw scores) whilst that for the 414 hearing-impaired children was 12·3 years (12·6 using raw scores).

The mainstreamed children in the present sample achieved mean maths ages of 14·0 years (14·4 using raw scores). Analyses of variance showed that their performance was significantly higher than that for the special school ($F = 33{\cdot}05$; $P < 0{\cdot}00000$) and partially hearing unit children ($F = 16{\cdot}78$; $P < 0{\cdot}00006$) and significantly lower than that for the ordinary hearing children ($F = 24{\cdot}86$; $P < 0{\cdot}00000$). Adding all hearing-impaired groups together, the mean maths age becomes 12·8 (13·1 using raw scores).

The mean hearing loss of the mainstreamed group was 48 dB. Breakdown into hearing loss bands is given in Table 1 together with a similar breakdown for children in special schools and units. Not surprisingly, most of the children in ordinary schools suffered only mild hearing losses in comparison with the other groups.

TABLE 1

PERCENTAGES OF CHILDREN IN DIFFERENT SCHOOL ENVIRONMENTS

With hearing losses between:	Special schools	PHU	Mainstream
30-60 dB	6	36	81
60-90 dB	31	38	17
90-120 dB	63	26	1

Sex differences in mathematical ability

In the original study significant sex differences—favouring boys—were found in the hearing sample, but not in the hearing-impaired one. In the present sample, there were 54 boys who achieved a mean maths age of 14·3 years and 81 girls who scored a mean 13·9 years. This difference is not statistically significant ($F = 0{\cdot}60$; $P < 0{\cdot}4$). When the three samples of hearing-impaired children are combined, they yield mean maths ages for the boys of 12·9 and for females of 12·7. This difference also is non-significant ($F = 0{\cdot}26$; $P < 0{\cdot}6$).

Although our first study revealed no significant differences between deaf boys and girls in terms of average levels of attainment, it did reveal different patterns of correlation between hearing loss and mathematical attainment. Pearson Product-Moment Correlation between hearing loss and attainment for girls ($r = -0{\cdot}29$; $P < 0{\cdot}002$) was significantly higher than that for boys ($r = -0{\cdot}04$; NS), using the Fischer z transform test ($P < 0{\cdot}007$). This pattern of results remains when we exclude children with losses of less than 30 dB.

In the mainstreamed sample, the correlation between hearing loss and mathematical attainment for boys was −0·14 and for girls −0·01. Here, neither correlation is significant but when we combine the scores of all three hearing-impaired groups, the correlation between hearing loss and achievement for all boys rises to −0·21 and that for all girls to −0·34 (yielding an overall correlation of −0·26; $P<0{\cdot}001$). Both correlations are significant ($P<0{\cdot}0005$ and 0·0000, respectively) but the difference between them is not ($P<0{\cdot}11$, two-tailed). Thus, the inclusion of the mainstreamed group removes the differential correlation between sex, hearing loss and mathematical achievement.

Relationships between school placement, sex, hearing loss and mathematical attainment

To explore the central question underlying this investigation—do children from different school environments perform differently when the effects of hearing loss are partialled out—multiple regression techniques were used including three predictor variables (placement, sex, and hearing loss) and one correlator variable (maths age).

The main results of the analysis are outlined in Table 2 below. The procedure involves two stages. In the first, different combinations of the three predictor variables are tested to determine which combinations—or ‘ models ’—have predictive power over maths age. In the second stage, each of the three models from which a different variable has been ‘ dropped ’ is tested, in turn, against that containing all three. This enables the identification of those variables that exert significant predictive power beyond that given by the other two. (This procedure was duplicated to include an interaction variable between sex and placement yielding similar results to those below.)

TABLE 2

Placement, Sex and Hearing Loss as Determinants of Maths Achievement

Model	Variables			multiple R	R square	F ratio	P
	Pl	Sex	HL				
1	1	2	3	0·28590	0·08174	11·906	0·0000
2		2	3	0·27006	0·07293	21·122	0·0000
3	1		3	0·27782	0·07718	14·943	0·0000
4	1	2		0·25646	0·06577	12·578	0·0000

Comparing Models

Models	Pl	Sex	HL	R squares	difference	F ratio	P
1	1	2	3	0·08174			
2		2	3	0·07293	0·00881	2·56562	0·0778
1	1	2	3	0·08174			
3	1		3	0·07718	0·00456	2·65422	0·1039
1	1	2	3	0·08174			
4	1	2		0·06577	0·01597	9·30342	0·0024

From Table 2 we can see that school background, sex, and hearing loss combined (model 1) produce a multiple correlation coefficient of 0·29. Although significant this accounts for only 8 per cent of the variance in maths scores, enabling us to make only very weak predictions about a child’s likely proficiency in maths age given only this information. Thus, the major determinants of mathematical ability in hearing-impaired children must lie outside the factors explored here.

Table 2 also reveals that all models successfully predict maths age. The problem is, of course, that although placement and hearing loss can both predict maths age, they also covary to a very large extent. But a comparison of the various models indicates that it is hearing loss which is the major contributor with school placement just failing to reach significance. Thus, comparing model 1 with model 2 (placement dropped) shows that losing information about placement does not have a significant effect on predictions about attainment. Conversely, the comparison between models 1 and 4 shows that dropping hearing loss as a factor leads to a significant drop in predictive power. Sex (dropped in model 3) also has no significant effect on predictive power.

The differences between the three groups of children can also be investigated by comparing the performances of children within the hearing loss bands that are relatively numerous in each group, i.e., between 50 and 70 dB. In the three groups—34 mainstreamed children, 44 in units, and 27 in schools—the average maths ages were 13·6, 13·3, and 13·1 respectively. Differences between these groups are not significant. Hearing loss thus remains as the major—if very weak—predictor of maths achievement.

Non-linear correlation between hearing loss and maths achievement

One possible explanation for the low correlation between maths ages and hearing loss is that we have adopted the wrong model. It is possible that the relationship is non-linear. In Table 3 we list a breakdown of average maths ages by different hearing loss bands across the various school placement groups. We also include the range in scores for each hearing loss band.

TABLE 3

AVERAGE MATHS AGES ACROSS HEARING LOSS BANDS AND SCHOOL ENVIRONMENTS (N in brackets)

HL (dB)	Mainstream	Units	Schools	Overall	Range
30-40	14·5 (48)	13·1 (7)	13·3 (2)	14·3 (57)	9·0-20·9
40-50	13·8 (40)	12·7 (21)	11·2 (2)	13·3 (63)	8·0-19·1
50-60	13·9 (22)	13·0 (20)	13·0 (12)	13·3 (54)	8·5-19·9
60-70	12·9 (12)	13·7 (24)	13·3 (15)	13·4 (51)	8·5-18·3
70-80	14·9 (7)	13·5 (13)	12·8 (22)	13·4 (42)	8·5-19·1
80-90	13·8 (4)	12·0 (14)	12·7 (46)	12·6 (64)	6·1-19·3
90-100	13·2 (1)	12·0 (19)	11·9 (65)	12·0 (85)	6·8-19·1
100-110	14·0 (1)	12·1 (10)	11·9 (71)	12·0 (82)	7·1-20·9
110-120	— (0)	12·3 (6)	11·1 (36)	11·3 (42)	5·8-18·7

No formal analysis is offered, but visual inspection of these relationships provides no evidence that any pattern other than a weak linear one would better fit the relationship between hearing loss and maths age. However, an analysis of the *range* and the *variance* of scores within each hearing loss band gives some insight into where there is likely to be most variability. Using Spearman Rank Correlation procedures, we can calculate that with increasing hearing loss there is a significant increase in variance ($r = 0{\cdot}83$; $P < 0{\cdot}01$). Thus children in the deafer groups produce more variable maths scores than those in the less deaf ones. The correlation between hearing loss and lowest scores achieved is significant ($r = 0{\cdot}82$; $P < 0{\cdot}01$). Thus, the lowest scores achieved within each hearing loss band vary systematically with hearing loss. However, there is no significant correlation between hearing loss and the highest scores achieved ($r = 0{\cdot}25$; NS). This suggests that although there are some very deaf children who are extremely proficient in maths (the 15 per cent of special education children who scored at or above their chronological age are no less deaf than the others: Wood *et al.*, 1983a), hearing loss interacts with other factors both to increase variability of attainment and to engender lower levels of achievement amongst the less able children. Such interactions could explain, in part, why the deaf sample produce lower mean scores.

Aetiology

One factor which we did not consider in the original study was that of aetiology or cause of deafness. Work in the United States has shown that cause of deafness is a major determinant of many educational abilities in deaf children (e.g., Jensema, 1975). For example, children who become deaf through maternal rubella score, on average, below children whose deafness is genetically linked on measures of language, mathematics, and literacy. In the present study, we attempted to solicit information about cause of deafness, and also age at onset of deafness. As in other studies, many children were reported as ' cause unknown '. Such cases (excluding 11 children for whom no aetiological information was given) comprised 46 per cent of the sample, which is similar to frequencies reported elsewhere (e.g., 42 per cent in a survey of deafness in the European Community: CEC, 1979).

In Table 4 we provide a breakdown of the sample of hearing-impaired children in ordinary schools into a number of aetiological categories. It should be noted that these classifications are derived from the categories specified by LEAs—we did not offer a formal classification system for their use.

TABLE 4

AGE AT ONSET OF DEAFNESS AND AETIOLOGY

	Number of Children	Average Hearing Loss (over 5 freq.)	Average Maths Age
Age at onset:			
Pre-lingual:	65	48·0	14·1
Post-lingual:	25	44·0	14·3
Unknown:	45	49·7	13·8
Cause of deafness:			
" Congenital ":	8	39·4	15·7
Known at birth: (e.g., jaundice, anoxia)	10	50·5	15·7
Rubella:	3	64·0	15·0
Familial:	8	47·6	14·8
" Hereditary/genetic ":	10	43·3	13·8
Ear infections/conductive:	16	39·4	13·7
Cause unknown:	57	49·4	13·6
No data:	11	44·0	12·7
Other causes:	12	58·7	14·7

Although the numbers of children in each group are too small to permit statistical analysis, it is noteworthy that the range of mean scores across groups (12·7 years to 15·7 years) roughly parallels the range of 12·1 to 15·5 years which were the means for children in special schools and ordinary hearing children.

One final feature of Table 4 is worth noting. The hereditary and familial groups represent 27 per cent of the known causes of deafness in this sample (and this does not include any of the 'congenitally deaf' children—12 per cent—who might have been meant to fall into this category), whilst the rubella group represents only 4 per cent. A recent EEC survey gave rubella as the largest single cause for children for whom cause was known. Sixteen per cent of the cause known group were rubella and

9 per cent genetic (as stated above, 42 per cent were cause unknown). Although incidence of deafness due to rubella fluctuates from year to year, this cannot account for the high relative incidence of the hereditary/familial groups in our sample. The clear implication of these figures is that special schools and units may be dealing with relatively more ' difficult ' children, such as rubella cases, and have a smaller proportion of genetically deaf children. This underlines the point made above—school placement and hearing loss are not the only factors contributing to the variance in mathematical achievements, and at least part of the reason for the superior performance of the mainstreamed children may stem not from degree of deafness or quality of teaching but differential problems associated with aetiology of the handicap.

DISCUSSION

Hearing loss, educational environment, and mathematical achievement

The main aim of this investigation was to determine whether or not the type of school attended by hearing-impaired children relates to their levels of achievement on a mathematics test when the effects of hearing loss are partialled out. The results show that it does not. Although degree of hearing loss correlates with mathematical achievement, it accounts for only a very limited amount of variance in the sample. Knowing how deaf a child is, what sex it is and what type of school it attends permits only weak predictive power over his or her likely levels of achievement. The major determinants of mathematical performance lie outside the factors explored in this study. Our results indicate that future attempts to study the basis of non-linguistic achievements in deaf children must look beyond these three factors if more than a trivial amount of variance is to be accounted for and if a fuller understanding of both the problems and abilities of deaf children is to be achieved.

Had the performance of the mainstreamed hearing-impaired sample more closely approximated those of hearing peers and been significantly better than those of special school children when the effect of hearing loss had been partialled out, then future research and educational innovation might have been directed at studies of the curriculum and at teaching methods being adopted in the different school environments. The fact that no additional variance in scores could be accounted for by placement suggests such a strategy would not prove productive. However, such studies might well help to explain variations in achievement *within* the different settings. In our previous research into the linguistic abilities of hearing-impaired children (e.g., Breslaw *et al.*, 1981; Howarth *et al.*, 1981; Wood *et al.*, 1982) we have reported significant differences in the performances of children with similar degrees of hearing loss and non-verbal intelligence drawn from different schools and classrooms. Such data indicate that school experiences and/or home environment exert an influence on deaf children's linguistic abilities.

Coupled with intervention studies demonstrating predictable short-term effects on children's use of language caused by differences in teaching style (e.g., Wood and Wood, 1984), such results suggest that the study of differential educational experiences might well increase our understanding of the basis of achievement in the deaf. Research in the United States leads to similar conclusions (Suppes, 1974; Jensema, 1975; Quigley and Kretschmer, 1982). The relatively small numbers of children drawn from any one school militated against inclusion of schools as a factor in our current analyses. What we can say, however, is that there is an extremely wide range in average maths ages across schools for the deaf in our sample. As one might expect, children from the two schools in the country which select for academic potential did produce extremely high results—averaging at 17·6 years of maths age. Across non-selective schools average scores ranged from 9·1 to 14·7 years. The nine severely and profoundly deaf children in the ' best ' non-selective school thus did better than the average for mainstreamed children. Though clearly not conclusive, such marked differences across schools in the same category suggest that more intensive studies might profitably be aimed at variation not across forms of provision but within each type of school environment.

Another cautionary note must be inserted about the relationships between educational environment and achievement. The test employed, though covering a range of mathematical abilities (including basic operations such as multiplication and division, algebra and interpretation of graphs), was chosen because it made few linguistic demands. Clearly, were a test demanding the application of mathematical reasoning to verbally stated problems to be employed, the relationships between degree of deafness and school environment might well be different. However, in relation to the development of basic numerical and computational skills, the finding remains that school attended, degree of hearing loss and, by implication, wide variations in linguistic ability, are not major determinants of achievement.

Sex differences in mathematical achievement

The results of the mainstreamed sample paralleled those of the special school and unit ones in the finding that sex differences in mathematical attainment are attenuated by deafness. Either deafness obscures the usual relationship between sex and mathematical performance by introducing additional sources of variance which mask that associated with sex, or the lack of difference for the deaf points to a social-communicative basis for such differences in hearing children (gender) rather than biological sex *per se*.

Previous studies have found that hearing loss (Wood *et al.*, 1983a) and linguistic ability (Norden, 1975) are correlated with maths performance in deaf girls but not boys. This might imply that such differential associations are bound up with differences in neurological organisation. However, such an hypothesis is not consistent with the results of the mainstreamed group. When their scores were added to those of the other samples, the differential associations between sex, hearing loss and performance were rendered non-significant. The reasons for this effect are not clear, but they do question our earlier suggestion that differences in neurological organisation are involved. Until more detailed investigations of the interactions between placement and other variables discussed below are explored, the reason for the differential relationships between hearing loss, sex and mathematical attainment remain obscure.

Other possible determinants

The range of information about children sought in this study was limited in the hope that relatively low (but still substantial) demands on the time of education authorities, teachers and children might encourage as many as possible to participate. This consideration, coupled with limited resources that prohibited personal contact with children and the collection of data about factors such as non-verbal intelligence, constrained the quantity and quality of information available.

However, the data on aetiological factors, though limited, points to their importance in the determination of performance. The nature of the interactions between aetiology, intelligence, sex, family background, learning difficulties, emotional problems and, as our study suggests, school attended, are extremely complex (Wood, in press). To date, no adequate theoretical framework exists to explain the interactions between this range of factors and, in consequence, our ability to evaluate educational provision and children's progress is limited. The probability that aetiology (and sex) influence decisions about school placement (albeit, perhaps, indirectly, through an association with rates of preschool development) further complicates attempts to compare school environments. Indeed, it could be argued that since special schools probably provide education for a disproportionate number of children with the most 'difficult' aetiologies, the absence of any differences associated with school environment is an indication of their success. Had environmental differences been found, with special schools, say, doing 'less well' than others, we would certainly have had to entertain the hypothesis that such differences could be attributable to aetiological factors.

In his book, *The Deaf School Child*, Conrad (1979) urges educationalists and researchers to acknowledge the importance of intelligence and degree of hearing loss in evaluating and studying the abilities of deaf children. To this list we would add aetiology and school factors.

Language and mathematical reasoning

In our previous theoretical considerations of the relationships between hearing loss, language, and mathematical reasoning, we have argued that studies of deaf children show that language in terms of 'inner speech' is not a *sine qua non* for computational skill and rational thought. The study of mainstreamed hearing-impaired children adds further credibility to this hypothesis. In this study, the attainments of a large sample of children with hearing losses ranging from 30 dB to 120 dB were only weakly related to degree of impairment. Similar studies of linguistic achievements (e.g., CEC, 1979) reveal far stronger correlations. We have studied children who are being educated alongside hearing children and who must, presumably, be coping with a normal or near-normal verbal environment. Their achievements, *vis-à-vis* the profoundly deaf children, suggest strongly that speech does *not* supply the code for computational reasoning.

This conclusion, however, leads us back to the paradox which, in part, inspired the current study. We still cannot explain why deaf children do less well than their hearing peers. One, still plausible, explanation is that a hearing loss, even quite a mild one, impedes the course of a child's learning because it interferes with the normal dynamics of communication and, hence, the rate of teaching and learning.

However, an examination of the relationships between *range* and *variance* in achievement and degree of hearing loss suggests that other factors would be worth exploring. Generally speaking, each hearing loss band contained children who scored at the ceiling of the test used. However, as hearing loss increased the lowest scores dropped still lower. There may, then, be some interaction between degree of deafness and other factors such as intelligence, aetiology and learning difficulties such that the variance in performance is increased with magnitude of hearing loss. There is a plausible case to be made for the hypothesis that the impact of any additional problems on the deafer child exerts a much greater effect on his levels of achievement than is the case for the less deaf child (Wood, in press), a process that we might term the 'multiplication of misery'. Such interactions are consistent with the finding of increasing variance in performance with degree of deafness and could, in part at least, explain why the hearing-impaired samples do less well than the hearing ones.

ACKNOWLEDGMENTS.—The research reported in this paper was supported by a grant from the Medical Research Council. We are grateful to all the Local Authorities, heads, teachers and children in the many schools who participated in this study.

REFERENCES

BRESLAW, P. I. B., GRIFFITHS, A. J., WOOD, D. J., and HOWARTH, C. I. (1981). The referential communication skills of deaf children from different educational environments. *J. child Psychol. Psychiat.*, 22, 269-282.

CEC: COMMISSION OF THE EUROPEAN COMMUNITIES (1979). *Childhood Deafness in the European Community*. Brussels-Luxemburg: ESSC-EEC-EAEC.

CONRAD, R. (1979). *The Deaf School Child.* London: Harper and Row.

HOWARTH, S. P., WOOD, D. J., GRIFFITHS, A. J., and HOWARTH, C. I. (1981). A comparative study of the reading lessons of deaf and hearing primary school children. *Br. J. educ. Psychol.*, 51, 156-162.

JENSEMA, C. J. (1975). *The Relationship between Academic Achievement and the Demographic Characteristics of Hearing-Impaired Children and Youth.* Series R, No. 2, Washington: Office of Demographic Studies, Gallaudet College.

NORDEN, K. (1975). Psychological studies of deaf adolescents. *Studia Psychologica et Pedagogica* Seria Altera (39). Lund: Gleerup.

QUIGLEY, S. P., and KRETSCHMER, R. E. (1982). *The Education of Deaf Children: Issues, Theory and Practice.* Baltimore: University Park Press.

SUPPES, P. (1974). Cognition in handicapped children. *Rev. educ. Res.*, 44, 165-176.

VERNON, P. E., and MILLER, K. M. (1976). *Graded Arithmetic-Mathematics Test.* Sevenoaks: Hodder and Stoughton.

WOOD, D. J., WOOD, H. A., GRIFFITHS, A. J., HOWARTH, S. P., and HOWARTH, C. I. (1982). The structure of conversations with 6- to 10-year-old deaf children. *J. child Psychol. Psychiat.*, 23, 295-308.

WOOD, D. J., WOOD, H. A., and HOWARTH, S. P. (1983a). Mathematical abilities of deaf school-leavers. *Br. J. dev. Psychol.*, 1, 67-73.

WOOD, D. J., WOOD, H. A., and HOWARTH, S. P. (1983b). Language, deafness and mathematical reasoning. In ROGERS, D. R., and SLOBODA, J. A. (Eds.), *The Acquisition of Symbolic Skills.* New York and London: Plenum Press.

WOOD, D. J. (in press). Social and educational adjustment in deaf children in relation to mental retardation. *Scientific Studies in Mental Retardation*, Royal Society of Medicine/Heinemann.

WOOD, H. A., and WOOD, D. J. (1984). An experimental investigation of the effects of five styles of teacher conversation on the language of hearing-impaired children. *J. child Psychol. Psychiat.*, 25, 45-62.

(*Manuscript received 22nd February*, 1984)

Br. J. educ. Psychol., 54, 265-275, 1984

THE EFFECT OF PICTURES ON THE ACQUISITION OF A SMALL VOCABULARY OF SIMILAR SIGHT-WORDS

BY ROBYN J. SAUNDERS AND ROBERT T. SOLMAN
(*University of New South Wales, Australia*)

SUMMARY. Two experiments were carried out to investigate the effects of pictures on learning to read a set of isolated, similar, common nouns. Five groups of kindergarten children took part in the first study. The children in four of the groups learnt each word using a picture of the thing represented by the word, and in two of these four pictorial groups the children had their attention drawn to the fact that the word and the picture represented the same thing. No mention was made of this relationship to the children in the other two groups, and within each of these association conditions one group of children was presented with the word 1 second prior to the presentation of the picture, and the other group received these inputs simultaneously. The fifth group of children acted as a control by learning the words in the absence of pictures. The results indicated that both the instructions to associate the word and the picture and the viewing of the word prior to the presentation of the picture failed to improve performance. Of more importance, however, was the observation that those children who learnt the words in the absence of pictorial information recognised more words on the critical post-learning trial than the children in the pictorial groups.

Four groups of kindergarten children took part in the second study. The children in two of the groups learnt the words with the aid of pictures, with one group seeing the picture 1 second prior to the word (the primed group) and the other group seeing the picture 1 second after the word (the non-primed group). The children in the remaining two groups were used as controls, with one group viewing a blank card 1 second prior to the word and the other group viewing the same card 1 second after the word. The results showed no difference in performance between the primed and non-primed conditions, but, consistent with the first study, the children who did not view pictures out-performed those who did. It was recommended that when teaching children to recognise visually presented words teachers abandon the general practice of using pictures as aids.

INTRODUCTION

THE general practice of using pictures as instructional aids when teaching young children to read words began in the early eighteenth century and, in spite of a growing body of literature which reports empirical findings indicating that this practice is at best of no benefit, it continues today. The first studies which questioned the use of pictures were reported by Miller (1937, 1938, 1939), and in 1967, Samuels noted the unpalatable fact that when differences in performance were obtained between groups of children taught using pictures and groups taught without their use the latter were superior. He accounted for this result by proposing that since the picture was the most salient input, its presence directed the child's attention from the word. This focal attention hypothesis has received considerable support (e.g., Braun, 1969; Duell, 1968; Beimiller, 1970; Singer *et al.*, 1973; Harzem *et al.*, 1976; Lang and Solman, 1979) but it does not explain results which fail to demonstrate improved performance when the child's attention is explicitly directed to the word (e.g., King and Muehl, 1965; Arnold, 1968; Kiraly and Furlong, 1974).

Lang and Solman (1979) extended the focal attention hypothesis in their examination of the effects of pictures on learning to read common nouns, by considering the repeated absence of a facilitatory effect of pictures from an information processing point of view. They agreed with Samuels (1967) that the picture probably directed the child's attention from the word, but they suggested that this might be because the processing capacity available to the child was not sufficient for him/her to process the two separate inputs at the same time. That is, in the absence of sufficient capacity to process the picture and the word simultaneously, the child was forced to choose one of these inputs to process first and in the majority of cases the more salient picture was selected. This hypothesis was tested by comparing the performance of children who viewed the word 1 second prior to viewing the picture, with the performance of children who viewed the two together. The spacing of the inputs in the former conditions allowed the children to focus attention first on the word and later on the picture, and a small advantage was observed for these children in all three of the experiments reported. However, these differences failed to reach significance except in a post-hoc experiment by conditions analysis of variance. The authors suggested that the absence of a more robust effect may have been a consequence of the fact that there were only small changes in performance in all experimental conditions, and these generally small changes were considered a consequence of the experimental procedure. Specifically, Lang and Solman administered only one acquisition and one test trial using a single list of 28 words, and many of the children who took part showed no evidence of learning.

Comparisons between conditions, where the words were viewed alone and conditions where they were viewed with pictures as aids, in the Lang and Solman (1979) experiments failed to reveal any advantage for the picture conditions. This absence of a facilitatory effect of pictures, while consistent with previous findings, may also have been due to the general absence of substantial learning effects, but the authors note that the structure of the word-list they used did not encourage the extraction of additional information from the pictures. Specifically, the 28 words they used were dissimilar, having different shapes and lengths and starting and ending in different letters, and the children may, therefore, have relied on these differences as cues for distinguishing the words. The effects of pictures on learning to read similar and dissimilar words have been investigated (King and Muehl, 1965; Hartley, 1970), but the definitions of similarity have differed and the results have been inconclusive. Hartley constructed a similar word list " by varying the initial consonant or the initial consonant cluster and holding the final elements of the word constant " (p. 107), and he concluded that pictures did not assist in learning these words. In contrast, King and Muehl considered that similar words were " approximately the same shape and contained some of the same sounds " (p. 164), and they concluded that pictures were useful for learning such words. If performance on similar and dissimilar words is to be measured, then the young child's perception of similarity should be used to define this dimension. A number of investigators (e.g., Marchbanks and Levin, 1965; Rayner, 1976; Rozin and Gleitman, 1977; Rayner and Posnansky, 1978) have shown that young children perceive as similar words which have the same first letter, the same last letter, similar shape and similar length.

In addition to the examination of the effects of sequencing the viewing of the word and the picture, the third study reported by Lang and Solman examined the effect on learning of explicit instructions to associate the picture and the word. The authors considered the results of this experiment to be promising, since the children who had their attention drawn to the fact that the word and the picture represented the same thing learnt more words than the children who were not instructed in this way. However, there was no control condition, where children viewed the word in the absence of pictorial information, used in this experiment, and as a consequence, no statement can be made concerning the general merits of the procedure. Also, the authors did not comment on whether they equalised the amount of verbalisation given to the children in the association and no-association conditions, and it might be reasonable to expect that more verbalisation was involved in instructing the children to associate the two inputs. Montare *et al.* (1977) have noted that children who are not presented with pictures frequently receive more verbal information than children in picture conditions, and they suggested that this results in better learning in the former case. By the same reasoning, therefore, the advantage demonstrated for the association conditions may have resulted from the more detailed instructions given.

The findings reported by Lang and Solman (1979) are potentially important. They suggest that if pictures are used as aids when teaching young children sight-words, they should be viewed after the word to avoid overloading the child's information processing system, and they should be explicitly associated with the word as this connection may not be made automatically. However, the points raised earlier in this introduction indicate that further empirical work is required before it can be confidently recommended that teachers adopt these procedures in the classroom. The first experiment reported here addresses this problem of replication.

EXPERIMENT 1

In the third experiment reported by Lang and Solman (1979), four groups of children learnt 28 common nouns with the aid of pictures. Two of these four groups were given explicit instructions to associate the picture and the word, and the other two groups were not instructed in this way. Within these association conditions, one group of children viewed the picture 1 second prior to viewing the word and the other group viewed the two inputs simultaneously.

These four experimental conditions were replicated in this study with the following modifications:

(1) Lang and Solman suggested that their studies failed to reveal substantial learning effects, because they administered only one acquisition and test trial using 28 words. To overcome this problem, four acquisition and test trials were administered using five words;

(2) Lang and Solman noted that because the words they used were physically dissimilar it was possible to distinguish among them without resorting to the information in the pictures. The five words used here were selected as being generally similar to each other;

(3) a control group of children who did not view words was included;

(4) the total amount of verbalisation by both the experimenter and the child was equalised for all five experimental conditions.

METHOD

Seventy children were randomly selected from the kindergarten classes of three New South Wales Government schools. Kindergarten is the first compulsory year of schooling in New South Wales, and all the children had been enrolled for approximately three months. Their ages ranged from 4·1 years to 6·1 years with an average value of 5·4 years. Equal numbers of boys and girls were selected, and a child was excluded from the sample if he/she had significant physical, social or emotional handicaps, obvious learning disabilities or, when given a preliminary test, was able to name any of the five visually presented stimulus words.

Seven boys and seven girls were randomly allocated to each of five experimental conditions. Care was taken during this process to control for learning outside the experiment by spreading group membership over the full data collection period, and to control for time-of-day dependent memory effects (e.g., Baddeley *et al.*, 1970) by spreading this membership evenly over the testing day (i.e., from 9.30 a.m. to 12.30 p.m.). Four of the groups of children viewed words and pictures and the fifth group viewed only words.

On each trial, two of the pictorial groups received instructions explicitly associating the word and the picture while the children in the other two groups did not receive these instructions. Within each of these association conditions, one group of children viewed the word 1 second prior to viewing the picture and the other group was presented with these two inputs together. The children in each of the five experimental groups were given the opportunity to view both the word and the picture for a total period of 5 seconds (a blank card replaced the picture for the children in the control group). That is, both the inputs were presented for a 5-second period, but in the case of the children in the spaced conditions, the word both arrived and was removed 1 second before the picture. The five stimulus words were randomly arranged prior to presentation on both acquisition and test trials, and correct responses were recorded and performance was measured as the number of correctly recognised words on the critical fourth trial.

The five stimulus words were selected on the basis of their ability to be easily represented by simple line drawings, their suitability for kindergarten children (based on teachers' judgments), their having the same first letter and the same last letter and their having similar length (it was necessary to vary the lengths of the words by one letter to satisfy the other criteria). The words selected were *shop*, *ship*, *step*, *stamp*, and *sheep*. They were all hand-printed in black ink on white cardboard sheets in lower case letters, and the realistic line drawings used as pictures were drawn on similar white cardboard also in black ink.

Each child attended a single experimental session which was conducted on an individual basis. The session was begun with a period of adjustment, during which the experimenter established rapport and administered a preliminary test. The preliminary test took the form of showing the child each word card for 20 seconds, and then asking if he/she could name the printed word. If the child named any of the words on this test he/she was excluded from the study. After the preliminary test, a child in one of the four picture conditions was familiarised with the pictures of the things named by the words, and a child in the control condition was familiarised with five pictures of other things.

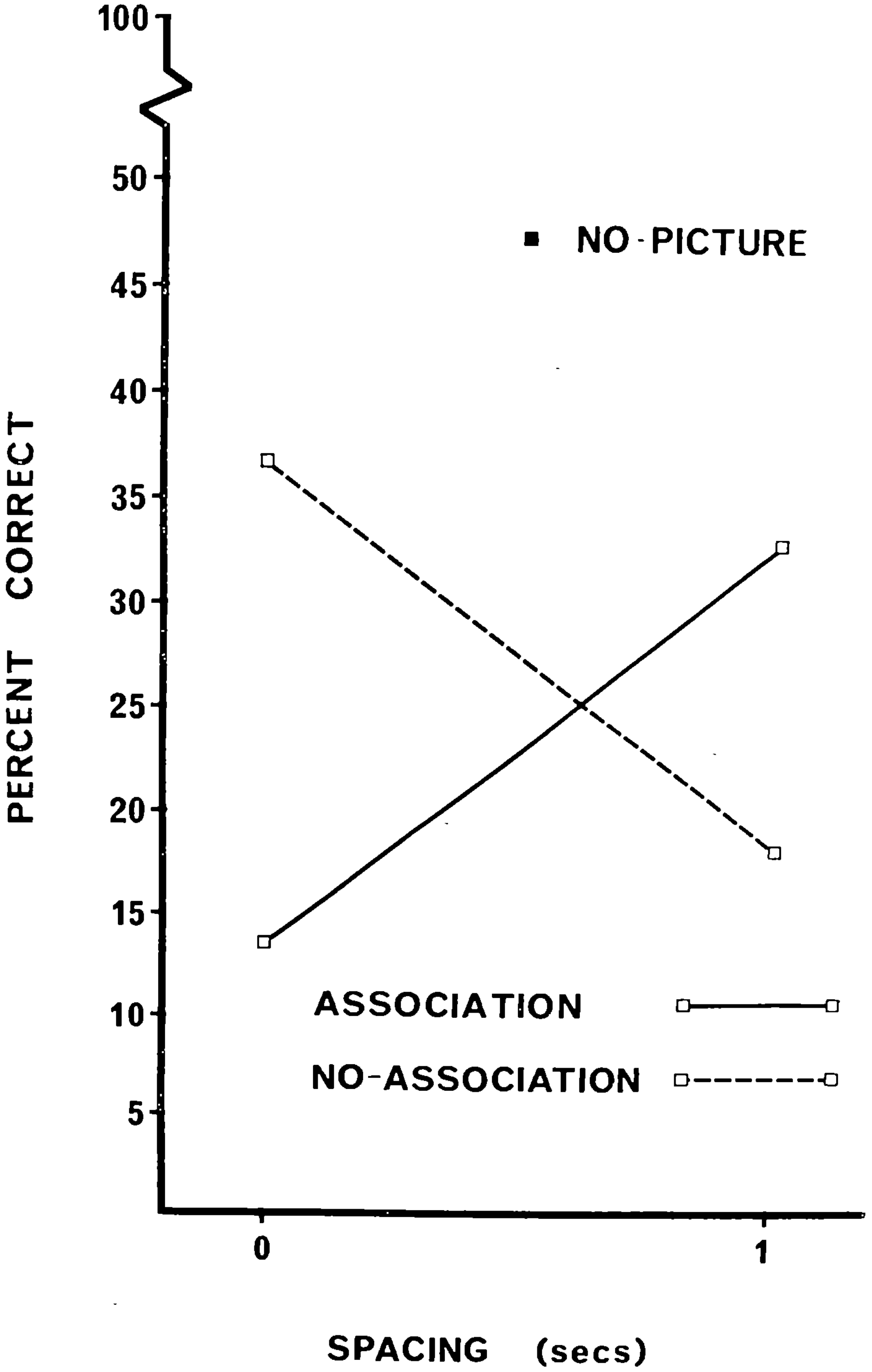
The first acquisition trial was introduced with the following words: " Now we shall begin to learn the words. If you can learn to read three of the five words I shall give you a gold star. I will give you another gold star for every other word you learn. Do you understand? " The instructions given to the child during acquisition depended on the experimental condition with, for example, the fact that the word and the picture represented the same thing being pointed out in the association conditions. However, the amount of verbalisation by both the experimenter and the child was equalised over all experimental conditions. The child was given feedback and an incorrect response was corrected during acquisition. Also, both the experimenter and the child repeated the name of the word on every trial. Upon completion of the first acquisition phase, the child was told that he/she would now be tested to see how many of the words he/she could recognise. The five word cards were then presented in a random order, one after the other, and the child was asked to name each word. No feedback or corrections were given during this phase. The acquisition and test phases were repeated four times, and the number of correctly named words during the critical fourth testing phase was recorded.

RESULTS

The percentage of correctly named words on the critical fourth testing trial is plotted for each of the five experimental conditions in Figure 1. Observation of these percentages suggests that, while both the explicit instructions to associate the word and the picture and the viewing of the word prior to the presentation of the picture failed

FIGURE 1

PERCENTAGE OF CORRECT RESPONSES OBTAINED ON THE CRITICAL TEST TRIAL FOR THE ASSOCIATION AND SPACED CONDITIONS AND FOR THE NO-PICTURE CONDITION



to improve performance, the effects of these two variables combined in a manner which resulted in the association instructions improving performance when the two inputs were spaced but depressing performance when they were simultaneous. Also, those children who learnt the words in the absence of pictures recognised more words on the critical testing trial than did the children in the four pictorial groups. An analysis of variance was carried out on the arcsin transforms (see Mosteller and Tukey, 1968) of the obtained data. This analysis contrasted the association with the no-association conditions, the spaced with the simultaneous conditions, and the no-picture control condition with the four-picture conditions. The results gave statistical support to the observations, in that while no significant effects were detected for association or for spacing ($F(1,65) = 0{\cdot}339$, $P > 0{\cdot}25$; $F(1,65) = 0{\cdot}010$, $P > 0{\cdot}25$ respectively), the interaction of these two variables reached significance ($F(1,65) = 6{\cdot}535$, $P < 0{\cdot}05$), and the difference between the no-picture and the picture conditions was significant ($F(1,65) = 6{\cdot}220$, $P < 0{\cdot}05$).

DISCUSSION

This study was designed to replicate the potentially important findings obtained by Lang and Solman (1979). These authors explained the fact that experimental investigations seldom indicated that pictures facilitated the learning of sight-words, by extending Samuels' (1967) focal attention hypothesis to incorporate the theory of limited capacity attention. They suggested that since the young child did not possess sufficient capacity to simultaneously process the word and the picture, he/she generally selected the more salient picture. This hypothesis was tested by comparing conditions where the word and picture were viewed together with conditions where the word was viewed for 1 second prior to the presentation of the picture, and a post-hoc analysis detected a small advantage for the latter, spaced conditions. They suggested that the absence of a more substantial effect of spacing was a consequence of their experimental procedure, and the procedure of this study was modified accordingly. Lang and Solman also found that children who had their attention drawn to the fact that the word and picture represented the same object, out-performed children who did not receive instructions to associate the two inputs.

The results of this study fail to demonstrate a consistent advantage for either the spaced or the associated conditions. The effects of these two variables did interact, however, and observation of Figure 1 suggests that this interaction might be described by noting that instructions to associate the word and the picture depress performance when they are presented together and improve performance when they are spaced. This pattern of differences is difficult to explain, but since there are no prior data or theoretical arguments which predict it, and since Scheffé post-hoc comparisons show that the difference between the two association conditions does not reach significance at either mode of presentation (i.e., $F(4,65) = 1{\cdot}230$, $P > 0{\cdot}25$; $F(4,65) = 0{\cdot}486$, $P > 0{\cdot}25$ respectively for simultaneous and spaced modes), a detailed account is not necessary at this stage.

The fact that children who viewed the word 1 second prior to the presentation of the picture did not recognise more words at the end of the acquisition period than did their classmates who viewed the word and the picture together indicates that the absence of a substantial effect of spacing in the Lang and Solman studies was not a consequence of their experimental procedure, and it suggests that the general absence of a facilitatory effect of pictures is not due to capacity limitations which prevent the simultaneous processing of the two inputs. We can conclude, therefore, that spacing the word and the picture with the former leading is at best of minimal benefit and cannot be recommended as a procedure to be adopted in the classroom.

The explicit instructions emphasising the fact that the word and the picture represented the same thing, which were given to children in two of the experimental groups, did not result in improved performance. Lang and Solman (1979) argued that, since young children experienced difficulty in forming imaginal associations (Wolff and Levin, 1972; Varley *et al.*, 1974), it could not be assumed that they automatically associated the word and the picture, and they suggested that the explicit instructions should be beneficial in forming this association. Their results supported this suggestion as the children who received the association instructions recognised more words than those who did not. However, there is no mention in their paper of whether care was taken to equalise the amount of instruction given to association and no-association groups, and it is likely that children in the former groups were instructed at greater length. It is possible, therefore, that the advantage for the association conditions resulted from a greater amount of verbalisation, not from differences in the nature of the instructions given. Care was taken to equalise the length of the instructions given to all groups in this experiment, and no consistent advantage was detected for the children in the associated groups. We cannot, therefore, concur with Lang and Solman's recommendation that the procedure of explicit instruction to associate the inputs be adopted in the classroom.

Lang and Solman did not include a no-picture control condition in their examination of the effects of spacing in combination with instructions to associate the word and picture. In this experiment, children who learnt the words in the absence of pictorial information recognised more of these words on the final trial than did those children who viewed both words and pictures. This is a persistent finding, and it is important in the present context since it shows that teachers cannot justify the use of pictures by referring to the facilitatory effects that Lang and Solman reported. That is, it is our opinion that the interpretation of experimental effects which result from different methods of presenting pictures is irrelevant unless at least one of these methods gives rise to a better level of performance than is demonstrated in a concurrent no-picture condition.

EXPERIMENT 2

The results of Experiment 1 showed that the use of pictures did not facilitate the learning of sight-words. Specifically, children who were not presented with pictures recognised more words than those who were, independent of the presentation method used and the instructions given. This suggests that we should recommend the abandonment of the practice of using pictures as aids. However, before we make this recommendation we will give the pictures " one more chance ", and reference to the literature on priming in semantic memory (e.g., Posner and Snyder, 1975; Neely, 1976, 1977) indicates a way in which this might be done.

Semantic priming is said to occur when experimental subjects are faster at identifying a test word when it is preceded by a related pre-test word, than when it is preceded by an unrelated word or is not preceded by any word. The related pre-test word is said to be a prime, and when it is presented activation spreads in semantic memory from its representation to the representations of semantically related words. This process of spreading activation concentrates attention on the activated network, and lowers thresholds and makes responses more readily available to words represented in the network. If we assume that the priming of a semantic network not only facilitates responses to represented words but also prepares the network for the incorporation of additional related material, then this procedure might also be used to aid learning. In other words, if it is reasonable to assume that the priming of a concept prepares the associative network for the integration of new related information, the picture of the object represented by the sight-word can be used as a prime.

The use of pictures as primes to aid the incorporation of related sight-words into the activated semantic network was examined in this study. That is, performance was compared for children who viewed the picture 1 second prior to the presentation of the word (the primed condition) with performance for children who viewed the word 1 second prior to the presentation of the picture (the non-primed condition).

METHOD

Fifty-six children were randomly selected from the kindergarten classes of two New South Wales Government schools. These schools were different from those used in the first experiment, and the children had been in attendance for approximately four months. Their ages ranged from 4·7 years to 6·3 years with an average value of 5·4 years, and they were selected and randomly allocated to four treatment groups using the procedures and applying the constraints detailed in the first study. The children in two of the four groups viewed words and pictures and those in the other two groups viewed only words. The two picture conditions differed, with the picture viewed 1 second prior to the presentation of the word when it was used as a prime, and with the word viewed 1 second prior to the presentation of the picture in the non-primed condition. The method used to present the two stimuli was the same as that used in the association-spaced condition in the first experiment. The primed and non-primed conditions were duplicated in the no-picture conditions, with a blank card substituted for the picture. Association instructions were, of course, not given in these latter conditions, but care was again taken to equalise the amount of verbalisation over all treatment groups.

The five stimulus words and the five stimulus pictures were those used in Experiment 1. Testing sessions were again conducted on a one-to-one basis, and followed the general procedure detailed for the first study. However, an additional acquisition trial was included and as a consequence, final performance was measured as the number of correctly recognised words on the critical fifth trial.

RESULTS

The percentage of correctly named words on the critical fifth testing trial is shown for each of the four treatment conditions in Table 1. Observation of these percentages suggests that the children who viewed the picture prior to the presentation of the word (i.e., the primed group), recognised more words on the final testing trial than their classmates who viewed the word prior to the presentation of the picture (i.e., the non-primed group). However, more obvious is the fact that the children who did not view the pictures as aids (i.e., the two no-picture groups), recognised more words than the children in both of the picture conditions. An analysis of variance was again carried out on the arcsin transforms of the obtained data. This analysis contrasted the primed picture condition with the non-primed picture condition, the two no-picture conditions with the two picture conditions, and the blank-card-first, " primed " no-picture condition with the blank-card-last, " non-primed " no-picture condition. The result of the comparison between the primed-picture condition and the non-primed picture condition did not provide statistical support for the observed advantage of the former ($F(1,52) = 1{\cdot}824$, $P > 0{\cdot}10$). In contrast, however, the advantage observed for the two no-picture conditions over the two picture conditions received statistical support ($F(1,52) = 8{\cdot}163$, $P < 0{\cdot}01$). No difference was detected between the two no-picture conditions ($F(1,52) = 0{\cdot}260$, $P > 0{\cdot}25$).

TABLE 1

PERCENTAGE OF CORRECT RESPONSES OBTAINED ON THE CRITICAL TEST TRIAL FOR THE PRIMED AND NON-PRIMED CONDITIONS AND FOR THE CORRESPONDING NO-PICTURE CONDITIONS

	Primed	Non-Primed
Picture	58·34	48·14
No-Picture	69·66	74·40

DISCUSSION

The most obvious result of this study is the fact that the children who did not view pictures as aids during acquisition recognised more sight words on the final testing trial than did their classmates who were presented with this additional information. This finding is consistent with the overall advantage obtained for the no-picture condition in the first study, and it again demonstrates the failure of pictures to facilitate learning. The observed advantage for the primed picture condition over the non-primed picture condition, although failing to reach statistical significance, might be considered worthy of further investigation. However, we would like to emphasise the point we made earlier in this regard, that is, the interpretation of experimental differences which result from different methods of presenting pictures is irrelevant unless at least one of the methods gives rise to better performance than can be demonstrated in a concurrent no-picture condition.

GENERAL DISCUSSION

The results obtained in the two studies reported in this paper fail to show that the viewing of pictorial representations of the things referred to by a set of sight-words aids the learning of these words. This is the case when children are given instructions which explicate the fact that the word and the picture represent the same thing, when the word is viewed 1 second prior to the presentation of the picture and when the picture is viewed as a prime 1 second prior to the presentation of the word. We recommend, therefore, that teachers abandon the practice of teaching young children sight-words by presenting the printed words in conjunction with related pictorial representations. (It should be noted that this recommendation applies only to the use of pictures when teaching isolated sight-words, and not to the use of illustrations associated with text. In the latter case the empirical evidence indicates that the pictorial information does, in fact, aid comprehension, e.g., Donald, 1979; Schallert, 1980.)

Why is it that the concurrent presentation of related pictures does not facilitate the learning of visually presented words? The procedure has intuitive validity, but when performance is compared for children taught using pictures with the performance of their classmates taught without their use, significant differences almost always favour the children in the latter no-picture conditions. The focal attention hypothesis (Samuels, 1967), with its straightforward suggestion that the pictures distract attention away from the words, cannot account for the absence of facilitation when procedures are used which direct the children's attention to the word and away from the picture. Also, the more elaborate limited capacity attention account (Lang and Solman, 1979), which considers that the child's processing capacity is limited to the extent that he/she cannot encode both the picture and the word at the same time, fails because it predicts a reliable advantage for children allowed to concentrate their attention first on the word and then on the picture and the experimental data do not support this prediction. Perhaps we should be looking for an explanation at a more sophisticated cognitive level?

When Lang and Solman proposed the use of explicit instructions associating the picture and the visually presented word, they referred to the work of Varley *et al.* (1974) and Wolff and Levin (1972) which showed that children's ability to form complex connected images develops with age. We have shown that the use of association instructions does not result in better learning of the sight-words, but the suggestion that the problem lies somewhere in the child's memory system is promising. Paivio (1971) considers that pictorial stimuli are automatically labelled on input and stored in separate imaginal and verbal systems. Thus, each picture that is capable of being labelled is stored twice. Words, on the other hand, are stored only in the verbal system unless they are " imaged ", in which case they are also stored in the imaginal system. Within this framework, the word and related picture always share a common verbal system, and it is within this system that they must be included in the same semantic network if the pictorial information is to aid in recalling the word. The experimental results indicate that this integrated network is not being formed, and we speculate that this is because its formation requires cognitive skill beyond that possessed by the young children who completed the learning tasks described in this article.

FOOTNOTE.—Requests for reprints should be addressed to Dr. Robert T. Solman, School of Education, University of New South Wales, PO Box 1, Kensington, NSW, Australia 2033.

REFERENCES

ARNOLD, R. D. (1968). Four methods of teaching word recognition to disabled readers. *Elem. School J.*, 68, 269-274.

BADDELEY, A. D., HATTER, J. E., SCOTT, D., and SNASHALL, A. (1970). Memory and time of day. *Br. J. Psychol.*, 22, 605-609.

BEIMILLER, A. (1970). The development of the use of graphic and contextual information as children learn to read. *Reading Res. Q.*, 6, 75-96.

BRAUN, C. (1969). Interest loading and modality effects on textual response acquisition. *Reading Res. Q.*, 4, 428-444.

DONALD, D. R. (1979). Effects of illustrations on early oral reading accuracy, strategies and comprehension. *Br. J. educ. Psychol.*, 49, 282-289.

DUELL, O. K. (1968). An analysis of prompting procedures for teaching a sight vocabulary. *Am. educ. Res. J.*, 5, 678-687.

HARTLEY, R. N. (1970). Effect of list type and cues on the learning of word lists. *Reading Res. Q.* 6, 97-121.

HARZEM, P., LEE, I., and MILES, T. R. (1976). The effects of pictures on learning to read. *Br. J. educ. Psychol.*, 46, 318-322.

KING, E. M., and MUEHL, S. (1965). Different cues as aids in beginning reading. *The Reading Teacher*, 19, 163-168.

KIRALY, J., and FURLONG, A. (1974). Teaching words to kindergarten children with picture, configuration and initial sound cues in a prompting procedure. *J. educ. Res.*, 67, 295-298.

LANG, R. J., and SOLMAN, R. T. (1979). Effect of pictures on learning to read common nouns. *Br. J. educ. Psychol.*, 49, 138-149.

MARCHBANKS, G., and LEVIN, H. (1965). Cues by which children recognize words. *J. educ. Psychol.*, 56, 57-61.

MILLER, W. A. (1937). The picture crutch in reading. *Elem. English Rev.*, 14, 263-264.

MILLER, W. A. (1938). Reading with and without pictures. *Elem. School J.*, 38, 676-682.

MILLER, W. A. (1939). What children see in pictures. *Elem. School J.*, 39, 280-288.

MONTARE, A., ELMAN, E., and COHEN, J. (1977). Words and pictures: critique of Samuel's findings *J. Reading Behav.* 9, 269-285.

MOSTELLER, F., and TUKEY, J. W. (1968). Data analysis, including statistics. In LINDZEY, G., and ARONSEN, E. (Eds.), *The Handbook of Social Psychology* (*2nd edition Vol. 2*). Reading, Mass.: Addison-Wesley.

NEELY, J. H. (1976). Semantic priming and retrieval from lexical memory: evidence for facilitatory and inhibitory processes. *Memory and Cognition*, 4, 648-654.

NEELY, J. H. (1977). Semantic priming and retrieval from lexical memory: roles of inhibitionless, spreading activation and limited capacity attention. *J. exper. Psychol.*, 106, 226-254.

PAIVIO, A. (1971). *Imagery and Verbal Processes*. New York: Holt, Rinehart and Winston.

POSNER, M. I., and SNYDER, C. R. (1975). Attention and cognitive control. In SOLSO, R. (Ed.), *Information Processing and Cognition: The Loyola Symposium*. Pontamac: Erlbaum Associates.

RAYNER, K. (1976). Developmental changes in word recognition strategies. *J. educ. Psychol.*, 68, 323-329.

RAYNER, K., and POSNANSKY, C. (1978). Learning to read visual cues to word recognition. In LESGOLD, A. M., PELLIGRINI, J. W., FAKKEMA, S. D., and GLASER, R. (Eds.), *Cognitive Psychology and Instruction*. New York: Plenum Publishing Corporation.

ROZIN, L., and GLEITMAN, L. R. (1977). The structure and acquisition of reading II. In REBER, A. S. and SCARBOROUGH, D. L. (Eds.), *Towards a Psychology of Reading*. Hillsdale, N.J.: Lawrence Erlbaum.

SAMUELS, S. J. (1967). Attentional processes in reading: the effects of pictures on the acquisition of reading responses. *J. educ. Psychol.*, 58, 337-342.

SCHALLERT, D. L. (1980). The role of illustrations in reading comprehension. In SPIRO, R. J., BRUCE, B. C., and BREWER, W. F. (Eds.), *Theoretical Issues in Reading Comprehension*. Hillsdale, N.J.: Erlbaum.

SINGER, H., SAMUELS, S. J., and SPIROFF, J. (1973). The effects of pictures and contextual conditions on learning responses to printed words. *Reading Res. Q.*, 9, 555-567.

VARLEY, W. H., LEVIN, J. R., SEVERSON, R. A., and WOLFF, P. (1974). Training imagery production in young children through motor involvement. *J. educ. Psychol.*, 66, 262-266.

WOLFF, P., and LEVIN, J. R. (1972). The role of overt activity in children's imagery production. *Child Dev.*, 43, 537-547.

(*Manuscript received 6th February*, 1984)

Br. J. educ. Psychol., 54, 276-283, 1984

THE TRANSFERABILITY OF TWO PICTORIAL SCIENTIFIC TASKS BETWEEN DIFFERENT SPATIAL DIMENSIONS

BY G. M. SEDDON, RIAZ H. TARIQ
(*University of East Anglia*)
AND J. DOS SANTOS VEIGA
(*University of Coimbra, Portugal*)

SUMMARY. Students in Cape Verde and Pakistan worked through different versions of two instructional programmes intended to teach the visualisation of reflection or rotation respectively in diagrams of three-dimensional structures. The three versions of the Reflections Programme were concerned respectively with reflections across a different Cartesian plane; the three versions of the Rotations Programme were concerned respectively with rotations about a different Cartesian axis. The results on a post-test indicated that all six programmes were successful on teaching the particular task in connection with the Cartesian plane or axis in which it had specialised. However, none of the programmes was successful in bringing about any learning to reflect across a different plane, or to rotate about a different axis.

INTRODUCTION

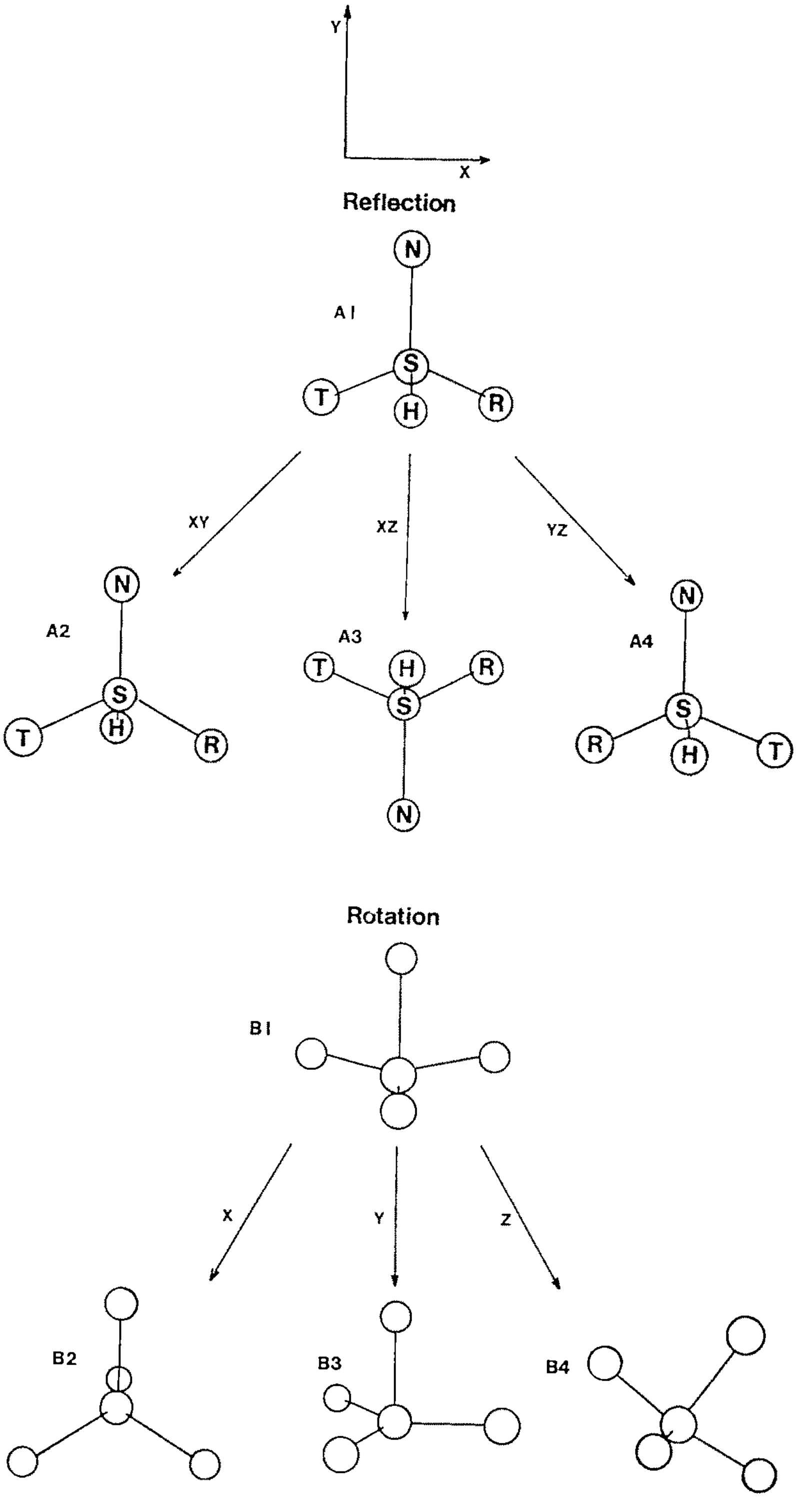
MANY scientific and technical subjects require students to visualise how diagrams should be drawn to represent the effects of performing spatial operations on three-dimensional structures. For example, as Figure 1 illustrates within the context of chemistry, students must understand that the diagrams A2, A3, and A4 represent the structure portrayed in diagram A1 after the structure has been reflected across each of the Cartesian planes, respectively. Similarly, it is essential for students to understand that diagrams B2, B3, and B4 are obtained by rotating the structure represented in diagram B1 about the X-, Y-, and Z-axes, respectively.

Similar tasks have been the topic of considerable theoretical and empirical investigation. Gibson (1966) has contributed notably to psychological theory, and Eliot and Hauptman (1981) have discussed the serious implications of previous work in relation to science education. More recently, Seddon *et al.* (1982) have demonstrated that alarmingly high proportions (i.e., 17-73 per cent) of very able science students in English sixth forms and a Portuguese university were unable to perform these tasks satisfactorily on diagrams such as those in Figure 1. The problems would seem to be even more severe in developing countries, where it is well known that students have difficulty in understanding spatial relationships in diagrams even without having to visualise the effects of a transformation (e.g., Miller, 1973; Walker, 1979; Petterson, 1982). The whole issue is therefore of world-wide educational importance.

While two successful methods of remedial teaching have been developed (Seddon *et al.*, 1984), there is a need to understand how they work in terms of present theories of understanding pictures (e.g., Gibson, 1966), which emphasise the role of the artistic devices or cues which are used to portray depth in diagrams. Thus for the diagrams being considered here, Figure 2 shows there are four depth cues which determine respectively the foreshortening of lines, the relative sizes of different parts of the structure, the representation of angles, and the extent to which different parts of the diagram overlap. The task of visualising the effects of a particular transformation amounts in essence to knowing how each of these depth cues changes as a result of that transformation.

FIGURE 1

An Illustration of the Effects of Performing Reflections and Rotations on Three-Dimensional Structures



C

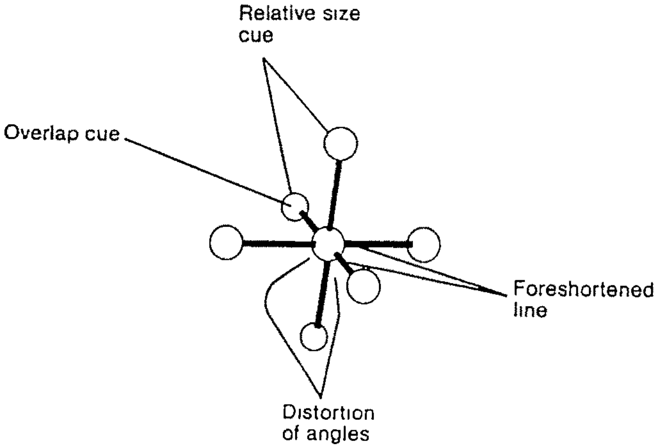
When a transformation involves motion in and out of the plane of the paper (i.e., a reflection across the XY-plane, or a rotation about the X- and Y-axes), Figure 1 shows that many depth cues must change. However, when the motion occurs either in or parallel to the plane of the paper (i.e., a reflection across the XZ- and YZ-planes, or a rotation about the Z-axis), there will generally be fewer changes in the depth cues. For example, Figure 1 shows that reflection across the XZ- and YZ-planes involves no change in the relative size, overlap or foreshortening cues. For rotation about the Z-axis, the example in Figure 1 indicates that the diagram of the rotated structure is obtained to a first approximation by simply rotating the diagram B1, as it stands, in the plane of the paper.

Since there appears to be a formal difference between the psychological nature of visualising transformations which involve motion in and out of the plane of the paper and those which do not, it is reasonable to suggest that there will be no or very little transfer between these two types of task. For example, students who have been taught to visualise reflections across the XY-plane may not simultaneously improve their performance on tasks which involve reflection across the XZ- and YZ-planes, and vice versa.

In contrast, reflections across the XZ- and YZ-planes are similar to each other in terms of these psychological processes, as are the rotations about the X- and Y-axes. Indeed, after rotating the diagrams through 90° in the plane of the paper, reflection across the YZ-plane is essentially the same as reflection across the XZ plane. Under similar conditions, rotation about the Y-axis is essentially the same as rotating about the X-axis. It may therefore be suggested that there will be transfer between these two types of reflection, and between these two types of rotation. For example, students who have been taught to visualise rotations about the X-axis, should also be able to perform better as regards rotations about the Y-axis.

FIGURE 2

AN ILLUSTRATION OF THE FOUR CUES USED TO PORTRAY DEPTH IN DIAGRAMS OF THREE-DIMENSIONAL STRUCTURES



(In the actual model being represented all the spheres are the same size; all the rods are the same length, and all the angles between adjacent rods are 90°)

On the other hand, there are also reasons for coming to the opposite conclusion as regards rotations about the X- and Y-axes. In particular, Seddon *et al.* (1982) showed that there are samples of students for whom the abilities to perform the reflections across the three Cartesian planes are factorially distinct. The same observation was also made for the three types of rotation. These results may imply that there can be no transfer between any of the different types of reflection, and between any of the different types of rotation. This possibility is also supported by other experiments by Seddon and his co-workers (i.e., Nicholson and Seddon, 1977; Santos Veiga *et al.*, 1981) who showed that the understanding of horizontal and vertical displacements were loaded only on their own orthogonal factors.

The present study aims to investigate the degree of transferability between the abilities to visualise the reflections, and between the abilities to visualise the rotations. The experiments are set in two developing countries, Cape Verde and Pakistan.

METHOD

The basic plan was to construct six instructional programmes each of which was designed to give students practice and direct training in visualising only one of the following types of transformation, i.e., reflection across the XY-, XZ- or YZ-planes, or rotations about the X-, Y- or Z-axes. Students were to be allocated to six groups, each group working through a different programme. Having worked through one of these programmes, the students were to be tested, as appropriate, on visualising all three reflections or all three rotations, and their performances compared with that of a randomly chosen control group which had not worked through any programme. A significantly better performance on any one of these three tests by an experimental group over that of the control group would imply that the students had learned to perform that task either by direct instruction or by transfer from the programme they had followed.

Subjects

The subjects were students aged 17-19 years following pre-university science courses in one college in Cape Verde and one college in Pakistan. All the students were well above average ability. They were always taught and tested in a second language, i.e., Portuguese in the case of Cape Verde, English in the case of Pakistan. In each college, 20 students selected at random from one academic year were assigned to each experimental or control group.

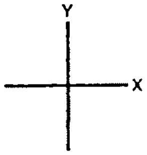
Post-tests

The post-tests for both reflections and rotations were modelled on those already used in previous work by Seddon *et al.* (1982). All the items were of the respective forms illustrated in Figure 3. Thus, the student was required to look at a diagram of a three-dimensional model, and select which of four other diagrams could represent the model after it had undergone the particular transformation specified in the question. In answering the question, the student was always able to refer to a diagram showing the orientation of the X- and Y-axes, and to a note explaining that the Z-axis was to be imagined as being perpendicular to the other two, and projecting in and out of the plane of the paper.

The diagrams are conventional ball-and-stick representations of molecules. All the balls are of the same diameter, and all the connecting rods of the same length. A total of 10 different configurations is used in the test, each configuration being one of those adopted by actual molecules. Details of the complete set of structures used are available from the authors.

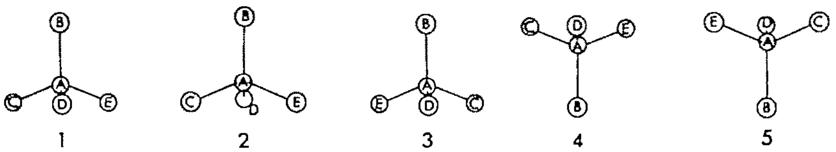
FIGURE 3

ENGLISH TRANSLATION OF SOME QUESTIONS USED IN THE REFLECTIONS AND ROTATIONS TESTS



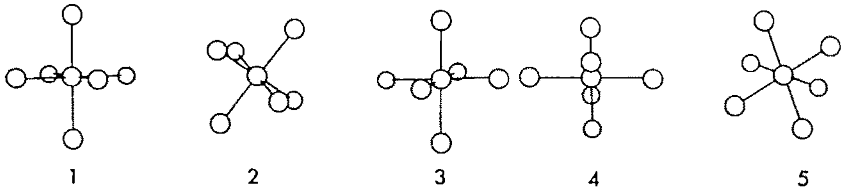
You need to refer to the axes in this diagram while answering the questions. The X- and Y- axes are drawn in the plane of the paper. You must imagine the Z-axis as being perpendicular to them coming out of the plane of the paper.

Reflection



If the structure in diagram 1 were rotated in the XY-plane, which of the diagrams 2–5 could represent the structure as seen after reflection?

Rotation



If the structure in diagram 1 were rotated about the Y-axis, which of the diagrams 2–5 could represent the structure as seen after rotation?

The diagrams were drawn using a computer program by Adamson and Cole (1969), to provide diagrams which are accurate in terms of perspective. The viewpoint for drawing diagram 1 was chosen arbitrarily. In drawing diagrams 2-5 for the Reflections Test, three were obtained by reflecting the model in diagram 1 across the three Cartesian planes, and the fourth by carrying out an inversion. For the Rotations Test, three of these diagrams were derived by rotating the model in diagram 1 about the X-, Y- and Z-axes respectively. The fourth diagram was obtained by rotating the structure about a randomly chosen pair of axes in sequence. For both tests, the order of presenting diagrams 2-5 in each question was determined randomly.

In administering the test, each question was printed as shown in Figure 2 in random order in a booklet, size A4. For the Cape Verde students, the test had been translated into Portuguese. Before the students began the test, the supervisor read out the simple procedural instructions which were also printed on the front cover of the booklet. These instructions also emphasised to the students that the diagrams represented three-dimensional structures. Then, after the supervisor had demonstrated the orientation of each axis, using a large model of labelled Cartesian axes, the students answered the questions under normal examination conditions, using a specially prepared answer card. There was no time limit, but all students finished in less than one hour.

Instructional programmes

The programmes were modelled on those found to teach successfully by Seddon *et al.* (1984). The basic idea is for the students to see a series of projected diagrams, each showing a structure, as seen from the same viewpoint, at a successive step, while undergoing a transformation. After seven such diagrams, the sequence is interrupted for the students to attempt a question of the type used in the post-test. This question is presented to the students in the appropriate language in a special booklet.

The question refers to the structure in the current sequence of diagrams, and asks the student to predict what the diagram would look like if the same transformation were continued beyond the stage represented by the last diagram. After a period of 15 seconds, the exposure of the remaining diagrams is resumed to reveal the correct answer to the question.

Each sequence within a programme dealt with a different molecular structure, and contained a total of 10 diagrams to represent equal intervals in the transformation. For the Rotations Programmes, these intervals corresponded to 10°, and were equal to those found to be the most effective in the experiment by Seddon *et al.* (1984). As with the post-test, the diagrams were drawn by computer, and were photographed with a high-quality camera to give 35 mm black-on-white slides each containing one diagram. In making and mounting the slides, great care was taken to ensure that, for each successive slide, the circles representing the central ball coincided exactly on the screen.

The slides were displayed with a conventional Carousel slide projector. Each slide was exposed for 10 seconds. The total time required to present the complete programmes was approximately 30 minutes. In all essential aspects, the conditions for presenting the different programmes were the same.

Administration

For each experiment, the randomisation was performed by allocating 20 students to a particular experimental group and 20 students to the control group. The instructional programmes and post-test were administered on successive days. Inevitably there were a few absentees for the post-test sessions.

RESULTS

For the sake of clarity, detailed results will be presented only for the Reflections experiment in Cape Verde (Table 1), and the Rotations experiment in Pakistan (Table 2). The detailed results for the other groups are essentially analogous, and lead to the same conclusions.

TABLE 1

SUMMARY OF THE SCORES OBTAINED ON THE REFLECTIONS TEST BY THE CAPE VERDE STUDENTS

Plane of Reflection		Plane of Reflection Taught			Control
Tested		XY (N = 20)	XZ (N = 20)	YZ (N = 17)	(N = 16)
XY	Mean	7·1**	4·2	1·4	4·1
	SD	2·7	4·2	2·4	2·8
XZ	Mean	3·1	8·6**	4·7	4·8
	SD	3·7	1·8	3·1	3·2
YZ	Mean	3·5	6·1	8·4*	5·5
	SD	3·4	4·2	1·6	3·8

Maximum score = 10 * $P < 0{\cdot}05$ ** $P < 0{\cdot}01$ (one-tail test)

TABLE 2

SUMMARY OF THE SCORES OBTAINED ON THE ROTATIONS TEST BY THE PAKISTANI STUDENTS

Axis of Rotation		Axis of Rotation Taught			Control
Tested		X (N = 20)	Y (N = 20)	Z (N = 20)	(N = 20)
X	Mean	4·6*	4·0	2·5	2·7
	SD	2·1	2·6	2·2	2·7
Y	Mean	2·8	5·8**	2·9	2·9
	SD	2·2	1·7	2·3	3·3
Z	Mean	3·8	4·5	5·9**	3·5
	SD	2·6	2·4	1·7	2·7

Maximum score = 10 * $P < 0{\cdot}05$ ** $P < 0{\cdot}01$ (one-tail test)

For some of the cells, it is clear that the distributions are skewed. The further analyses were therefore conducted using both non-parametric and parametric tests. Both methods lead to the same conclusions.

On subjecting the data for each post-test to a one-way analysis of variance, all six gave a significant value of F ($P < 0{\cdot}01$). The means of each experimental group were then compared with that of the corresponding control. The parametric analyses used Dunnett's t-test for this purpose. All the significant comparisons are indicated by the asterisks in Tables 1 and 2.

DISCUSSION

The results in Tables 1 and 2 indicate that all six programmes (i.e., reflections plus rotations) had a significant effect in improving the students' performance on the transformation which was actually practised in the programme. The increases in the mean scores for the Rotations Tests are actually quite small in numerical terms, and are considerably smaller than the corresponding increases found by Seddon *et al.* (1984) using essentially the same programmes with Nigerian students. Clearly, there is room for improvement in the development of the programmes for these students. It is very probable that, as Seddon *et al.* (1984) found, significantly better results would have been obtained using a lap-dissolve control unit to simulate motion more realistically in changing the slides. However, such devices are not readily available in these countries, and perhaps the most promising method for improvement is simply to provide additional sequences and examples to give the students more practice on the specific transformations.

The results showed no examples of positive transfer in either direction between any of the reflections, or between any of the rotations. There was actually one case of negative transfer in that students who had been taught to reflect across the XY-plane performed significantly worse than the control group on reflections across the YZ-plane.

The lack of transfer to or from reflections across the XY-plane, and to or from rotations about the Z-axis is entirely in agreement with the previous theoretical predictions. The results would seem to support the validity of using differences in the effects of changing the depth cues as a way of delineating psychological differences between the tasks.

In contrast, the observed lack of transfer between the XY- and XZ-reflections, and between the X- and Y-rotations does not support this approach. It is clear that the students were quite unable to see the similarities between these two types of reflections, and between these two types of rotations. The observed results are more in line with the conclusions to be drawn from the factor analyses by Seddon *et al.* (1982), i.e., that tasks within these two pairs were factorially distinct.

On the other hand, it must be borne in mind that these negative results have been obtained using only samples of 20 students for each treatment group. If larger samples had been used, it is possible that the null hypotheses could have been rejected with the same differences in the mean scores observed in the present experiment, and thereby have led to different conclusions. Moreover, this possibility cannot be discounted without repeating the experiments on a larger scale. However, if statistically significant differences between groups are observed using only very large samples, these effects are too small to be of practical interest; they are of only theoretical interest.

In contrast, the present experiment has been carried out using samples which are comparable in size to the classes which are taught in these institutions, and if the results are to be of practical relevance, the effects of the conclusions and recommendations which are made must be observable in this context. For these purposes, it must be concluded that there is no positive transfer between any of the reflections, and between any of the rotations. It is therefore essential that any instructional programme should give the students practice and training in all three reflections and all three rotations.

Acknowledgment—The authors wish to thank the British Council, Gulbenkian Foundation and Leverhulme Trust for travel grants to carry out this investigation.

REFERENCES

Adamson, P. G., and Cole, A. J. (1969). A simple method for drawing molecules using a digita plotter. *Acta Crystallographica A25*, 535-539.

Eliot, J., and Hauptman, A. (1981). Different dimensions of spatial ability. *Studies in Science Education*, 8, 45-66.

Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems.* Boston: Houghton Mifflin.

Miller, R. J. (1973). Cross-cultural research in the perception of pictorial materials. *Psychol. Bull.*, 80, 135-150.

Nicholson, J. R., and Seddon, G. M. (1977). The understanding of pictorial spatial relationships by Nigerian secondary school students. *J. cross-Cultural Psychol.*, 8, 381-400.

Petterson, R. (1982). Cultural differences in the perception of image and color in pictures. *Educ. Commun. Technol. J.*, 30, 43-53.

Dos Santos Veiga, J., Seddon, G. M., and Jusoh, I. (1981). Testing the understanding of pictorial spatial relationships using the Portuguese and Malay languages. *Educ. Psychol.*, 4, 305-318.

Seddon, G. M., Eniaiyeju, P. A., and Jusoh, I. (1984). The visualization of rotation in diagrams of three-dimensional structures. *Am. educ. Res. J.*, 21, 25-38.

Seddon, G. M., Tariq, R. H., and Dos Santos Veiga, J. (1982). The visualisation of spatial transformations in diagrams of three-dimensional structures. *Eur. J. sci. Educ.*, 4, 409-420.

Walker, D. A. (1979). *Understanding Pictures: a Study in the Design of Appropriate Visual Materials for Education in Developing Countries.* Massachusetts: University of Massachusetts, Centre for Instructional Education.

(*Manuscript received 16th October*, 1983)

Br. J. educ. Psychol., 54, 284-292, 1984

ACADEMIC SELF-CONCEPT: SOME DEVELOPMENTAL AND EMOTIONAL CORRELATES IN NINE-YEAR-OLD CHILDREN

BY JAMES W. CHAPMAN,
(*Massey University, New Zealand*)
PHIL A. SILVA, AND SHEILA M. WILLIAMS
(*Dunedin Multidisciplinary Research Unit, University of Otago, New Zealand*)

SUMMARY. Eight hundred 9-year-old children being studied by the Dunedin Multidisciplinary Health and Development Research Unit were assessed in terms of academic self-concept, intelligence, and school performance. A short form of the Student's Perception of Ability Scale (SPAS) was used to estimate academic self-concept. Intelligence was assessed by the Wechsler Intelligence Scale for Children–Revised (WISC–R), reading by the Burt Word Reading Test, spelling by a short form of the Dunedin Spelling Test, comprehension and expression by the respective subscales in the Illinois Test of Psycholinguistic Abilities (ITPA), and school achievement by parents' perceptions of current attainment levels. The results indicate that the SPAS Short Form yields means, standard deviations, and internal reliability estimates that are consistent with other studies. In addition, sex differences were noted, with girls reporting slightly higher self-perceptions of ability than boys. In terms of correlations between academic self-concept and school-related ability and performance, it was found that SPAS subject-specific subscale scores correlated moderately with corresponding subject area performance, but weakly or negligibly with WISC–R and ITPA Comprehension and Expression scores. Perceptions of ability, therefore, appear to be relatively independent of intelligence, and more related to success levels in school. It was also found that children identified as reading retarded had significantly lower self-perceptions of ability than normal readers. Similarly, depressed children indicated relatively negative self-concepts compared to past depressed and non-depressed children.

INTRODUCTION

BLOOM (1976) asserted that three major input components influence performance on learning tasks. These " inputs " are " cognitive entry behaviours ", " quality of instruction ", and " affective entry characteristics ". While considerable attention in educational psychology has focused on cognitive and instructional aspects of human learning, relatively little empirical effort has been applied to investigating the influence of affective factors in learning. Indeed, some writers (e.g., Wylie, 1974; Wells and Marwell, 1976) have noted that during the second quarter of this century the dominant behaviourist and functionalist psychologies left little room for such mentalistic elements as affective influences on behaviour.

More recently an upsurge of interest in self-referent constructs has become apparent. This interest is particularly evident in the areas of self-concept and self-efficacy (Covington and Beery, 1976; Wells and Marwell, 1976; Bandura, 1980), and causal attributions (Weiner, 1980). Research efforts here strongly suggest that self-perceptions and conceptions are important mediational influences which define for individuals the nature of their relationships with other people, the types of behaviours and tasks in which they will engage, the states of tension they will experience, and in turn, how individuals will perceive themselves (e.g., Kifer, 1975; Abramson *et al.*, 1978; Bandura *et al.*, 1980; Weiner, 1980). Clearly then, self-perceptions of competence and controllability can be viewed as causal agents which, through their defining roles in an individual's phenomenal field, are active constituents in the prediction of human behaviour.

In terms of self-referent constructs and achievement behaviours, an increasing number of studies have addressed the relationship between *general* self-concept and learning outcomes (Purkey, 1970; Uguroglu and Walberg, 1979; Revicki, 1982). However, relatively few studies have focused specifically on *academic* self-concept. Yet Shavelson *et al.* (1976) have argued that the more closely self-concept is linked with specific situations the closer is the relationship between self-concept and behaviour in the specific situation. Thus, academic self-concept should have a stronger relationship with school achievement than general self-concept—a commonsense belief that has received some empirical support (Boersma and Chapman, 1978; Uguroglo and Walberg, 1979; Burns, 1982).

One of the main reasons for the lack of research on facet-specific dimensions of self-concept has been the inadequate instrumentation. This inadequacy is reflected in both the lack of acceptable reliability and validity characteristics as well as the global nature of most self-concept scales (Wylie, 1974). The recent development of the Student's Perception of Ability Scale (SPAS: Boersma and Chapman, 1977; Boersma *et al.*, 1979), however, has provided a measure of school specific perceptions of ability that has satisfactory psychometric characteristics. Prior to that time, the only frequently used measure of *academic* self-concept noted in the literature was the Self-Concept of Ability Scale, developed by Brookover and associates during the early 1960s for use with high school students (Brookover *et al.*, 1965).

The development of the SPAS allows the investigation of academic self-concept and its relationship to a range of school-related cognitive and social-emotional variables. In the present study the focus is on academic self-concept, measured intelligence, achievement, and emotional status in a large sample that is part of a multidisciplinary longitudinal study of child health and development. In addition, academic self-concept is investigated in a sub-sample of depressed children, and a sub-sample of children with specific reading retardation.

METHOD

Sample

The sample comprised 800 children (415 boys and 385 girls) aged 9 years, who are currently being studied by the Dunedin Multidisciplinary Health and Development Research Unit. This is a longitudinal study involving some 1,037 children born at one hospital in Dunedin, New Zealand during the 12-month period between 1st April, 1972 and 31st March, 1973 (McGee and Silva, 1982). Of these 1,037 children currently enrolled in the study, 955 were assessed on a number of physical, medical, and psychological measures around the time of their ninth birthday (87 per cent within one month). The 800 children in the present sample represent those who completed the measure of academic self-concept, with the 155 children who did not complete that particular measure being deleted from the analyses.

Measures

Academic self-concept was measured with the Student's Perception of Ability Scale (SPAS: Boersma and Chapman, 1977). The SPAS contains 70 declarative " Yes-No " items relating to feelings and attitudes about school performance in five basic academic areas (reading, spelling, language arts, arithmetic, and penmanship), and also to school in general. The items contribute to six subscales, derived through factor analysis, which include Perception of General Ability, Perception of Arithmetic Ability, General School Satisfaction, Perception of Reading and Spelling Ability, Perception of Penmanship and Neatness (each of which contains 12 items), and Confidence in Academic Ability (10 items). The SPAS technical data are reported elsewhere (Boersma and Chapman, 1979; Boersma *et al.*, 1979).

In the present study a shortened version of SPAS was used, consisting of 35 items.

These were selected on the basis of their subscale factor loadings (Boersma and Chapman, 1979): thus, half the number of items in each subscale, those with the highest loadings, were included in the shortened form. Accordingly, five of the subscales had six items, whereas the sixth subscale (Confidence) had five items. Examples of items included in the short form included " I am unhappy with how I read ", " I am good at spelling ", " I like arithmetic ", and " I have problems printing neatly ".

Technical data for the SPAS Short Form (Chapman *et al.*, 1983) reveal that the mean Full Scale score for 9-year-olds is 23·12, which is approximately half the mean of 45·47 reported in the SPAS Manual (Boersma and Chapman, 1979) for Grade 4 (9-year-old) children. Similarly, the Short Form Full Scale standard deviation of 5·71 is close to half the Grade 4 standard deviation of 12·10. Internal consistency estimates (Cronbach's alpha) are acceptable for a scale reduced by half, with the Full Scale alpha being 0·81, and the Subscale alphas ranging from 0·81 to 0·57.

Measures of intellectual ability, academic achievement, and general social-emotional behaviour were obtained using a number of instruments. Intelligence estimates were based on scores on the Wechsler Intelligence Scale for Children–Revised (Wechsler, 1974). Reading levels were established by the Burt Reading Test (Scottish Council for Educational Research, 1976) while spelling ability was assessed by a short form of the Dunedin Reading Test (Silva *et al.*, 1983). Verbal comprehension and expression were measured by the respective subtests of the Illinois Test of Psycholinguistic Abilities (ITPA: Kirk *et al.*, 1968). Finally, emotional status was determined by the Child Affective State Questionnaire (Kashani *et al.*, in press) and depression with the KIDDIE-SADS-E scale (Orvaschel *et al.*, 1982).

Procedure

All data were collected on or about each child's ninth birthday. The questionnaires and tests were completed within one month of their ninth birthday by 87 per cent of the sample. Most children and parents in the sample (96·5 per cent) participated in this phase of the study at the research centre in Dunedin city. However, 3·5 per cent of children enrolled in the study but residing outside Dunedin completed the tests and questionnaires in other towns or cities, usually under the supervision of a psychologist from the New Zealand Psychological Service.

For those children participating in the study at the research unit, the SPAS was presented to each child by a psychometrist and, following an explanation of the instructions, was self-completed. IQ, reading, spelling, and ITPA tests were administered by psychometrists following completion of the SPAS. Mothers answered the Child Affective State and school achievement questionnaires independently at the same time. Finally, the K-SADS-E was administered by a psychiatrist to a sub-sample of 121 children.

RESULTS

Academic self-concept

Descriptive SPAS data as a function of sex are presented in Table 1. A statistically significant difference in mean Full Scale scores between boys and girls was evident. Girls tended to obtain slightly higher SPAS scores overall than boys. This finding is consistent with findings from the Canadian normative study (Boersma *et al.*, 1979). The difference between the sexes of 1·2 points, however, was well within the SEm value for the Short Form (2·51) and therefore of little practical importance. Small, but statistically significant, differences in scores between boys and girls were also observed on the School Satisfaction, Reading/Spelling, and Penmanship/Neatness subscales. In all cases, girls obtained higher mean subscale scores than boys. These differences were again consistent with the normative study.

TABLE 1

DESCRIPTIVE DATA FOR SPAS SCORES AS A FUNCTION OF SEX

	Boys (N = 415)		Girls (N = 385)		
	$\bar{x}$	SD	$\bar{x}$	SD	t^a
Full Scale	22·54	5·71	23·74	5·64	3·00*
General Ability	3·51	1·77	3·69	1·71	1·50
Arithmetic	4·61	1·65	4·49	1·75	1·03
School Satisfaction	3·75	1·59	4·06	1·41	2·90*
Reading/Spelling	3·99	1·81	4·48	1·70	3·94*
Penmanship/Neatness	4·73	1·74	5·04	1·55	2·63*
Confidence	1·94	1·42	1·98	1·47	0·39

^{a} df = 798 * P < 0·01

SPAS correlations with other variables

Product-moment correlations between SPAS Short Form scores and measures of intelligence, achievement, and emotional status are shown in Table 2. Coefficients greater than 0·11 were statistically significant at the P < 0·001 level. For the IQ data, Full Scale SPAS correlations with WISC-R scores were all low (0·27 to 0·12) suggesting that IQ was not meaningfully related to academic self-concept in the Dunedin sample. Of the SPAS subscales, General Ability showed the highest correlations, ranging from

TABLE 2

PRODUCT-MOMENT CORRELATIONS BETWEEN ACADEMIC SELF-CONCEPT, INTELLIGENCE, AND ACADEMIC PERFORMANCE

	Full Scale	General Ability	Arithmetic	School Satisfaction	Reading/ Spelling	Penmanship/ Neatness	Confidence
WISC–R							
Full IQ	0·262***	0·380***	0·251***	−0·068*	0·225***	0·055	0·013
Verbal IQ	0·259***	0·384***	0·243***	−0·052	0·237***	0 024	0·011
Performance IQ	0·199***	0·288***	0·196***	−0·069*	0·152***	0·079*	0·002
Information	0·239***	0·344***	0·225***	−0·082*	0·230***	0·028	0·034
Similarities	0·152***	0·278***	0·150***	−0·012	0·133***	−0·023	−0·033
Arithmetic	0·262***	0·301***	0·279***	−0·041	0·237***	0 060	0 023
Vocabulary	0·203***	0·330***	0·134***	−0·044	0·191***	0·042	0·011
Picture Completion	0·118***	0·193***	0·095**	−0·023	0·064	0·072*	−0·018
Block Design	0·140***	0·260***	0·181***	−0·126***	0·118***	0·018	−0·008
Object Assembly	0·123***	0·186***	0·136***	−0·031	0·073*	0·052	−0·016
Coding	0·200***	0·195***	0·156***	−0·001	0·183***	0·096**	0·037
Burt							
Reading	0·358***	0·372***	0·165***	−0·019	0·456***	0·075*	0·146***
DMHDS							
Spelling	0·365***	0·325***	0·136***	−0·007	0·486***	0·113***	0·172***
ITPA							
Comprehension	0·137***	0·252***	0·109**	−0·057	0·125***	0·061	−0·055
Expression	0·130***	0·102**	0·076*	0·077*	0·066*	0·029	0·031
School							
General	0·317***	0·341***	0·218***	−0·009	0·326***	0·050	0·136***
Reading	0·332***	0·346***	0·165***	−0·028	0·389***	0·079*	0·159***
Spelling	0·317***	0·274***	0·164***	0·000	0·399***	0·062	0·169***
Arithmetic	0·279***	0·276***	0·310***	−0·029	0·206***	0·043	0·133***
Writing	0·289***	0·297***	0·168***	0·018	0·296***	0 096**	0·093**
Printing	0·214***	0·174***	0·088**	−0·009	0·210***	0·183***	0·073*
Affective State	0·425***	0·408***	0·261***	0·108**	0·334***	0·276***	0·079*

* P < 0·05 ** P < 0·01 *** P < 0·001

0·38 (Verbal IQ) to 0·19 (Object Assembly). For SPAS Arithmetic, the highest correlation was with WISC–R Arithmetic (0·28). SPAS Reading/Spelling also had its highest correlation with WISC–R Arithmetic and also with WISC–R Verbal IQ, being 0·24 in both cases. The other SPAS subscales showed only weak to negligible correlations with WISC–R scores.

Results for the Burt Reading and Dunedin Spelling tests reveal that in both cases the highest correlations were with SPAS Reading/Spelling, with coefficients of 0·46 and 0·49 respectively. Moderate correlations were recorded for Burt Reading with the SPAS Full Scale (0·36) and General Ability (0·37), and for spelling with the SPAS Full Scale (0·37) and General Ability (0·33).

In terms of the ITPA, correlations were only weak to negligible for both Comprehension and Expression. Of these two scales, the highest correlation was only 0·25 for Comprehension with SPAS General Ability.

Correlations between the SPAS scales and parents' perceptions of their child's school progress revealed moderate to negligible correlations. The strongest relationship here appear to be for SPAS Reading/Spelling with Reading (0·39) and Spelling (0·40). SPAS Arithmetic correlated highest with Arithmetic (0·31) and SPAS General Ability highest with General School progress (0·34).

Finally, correlations between SPAS scores and the child's Affective State indicated a moderate overall relationship between academic self-concept and emotional status. The SPAS Full Scale correlation with the Child's Affective State was 0·42, whereas SPAS subscale correlations ranged from 0·41 (General Ability) to 0·08 (Confidence).

Group comparisons for selected subsamples of children were made in terms of specific reading retardation and depression. Specific reading retardation was defined by means of an adaptation of the Yule *et al.* (1974) regression formula, in which predicted reading score is calculated using the WISC–R Performance IQ score. This procedure resulted in 36 (32 boys and 4 girls) of the 800 children being identified as having average intelligence (Delay IQ $\bar{x} = 104{\cdot}2$, SD = 14·68; Remainder of sample IQ $\bar{x} = 105{\cdot}4$, SD = 16·38) but a reading delay of at least two years.

SPAS scores for the specific reading retardation group was compared with the remainder of the sample in Table 3. Here it can be seen that there was a significant difference in Full Scale scores between the two groups, with the specific reading retardation group reporting lower SPAS scores than the normal group: (Delay $\bar{x} = 20{\cdot}33$, SD = 5·26; Remainder of Sample $\bar{x} = 23{\cdot}25$; SD = 5·70; t = 3·01, df = 797, $P < 0{\cdot}01$). Significant subscale differences between the two groups were evident for General Ability ($P < 0{\cdot}05$) and Reading/Spelling ($P < 0{\cdot}001$).

TABLE 3

SPAS FULL AND SUBSCALE SCORES FOR NORMAL AND DELAYED READERS

	Delayed Readers (N = 36)		Normal Readers (N = 764)			
	$\bar{x}$	SD	$\bar{x}$	SD	t^a	P
Full Scale	20·33	5·26	23·25	5·70	3·01	0·003
General Ability	2·89	1·64	3·63	1·74	2·63	0·012
Arithmetic	4·36	1·74	4·56	1·70	0·68	0·499
School Satisfaction	3·97	1·36	3·90	1·52	0·32	0·748
Reading/Spelling	2·56	1·72	4·30	1·74	5·97	0·000
Penmanship/Neatness	4·64	1·64	4·89	1·66	0·91	0·368
Confidence	1·92	1·30	1·96	1·45	0·21	0·835

[a] df = 797

Finally, SPAS results were compared for groups of current depressed, past depressed, and non-depressed children. The currently depressed group consisted of 17 children (10 boys and 7 girls), whereas the past depressed group consisted of 23 children (8 boys and 15 girls). A comparison group of 76 non-depressed children (38 boys and 38 girls) was also identified following psychiatric interviews. The three groups were not significantly different from each other in terms of IQ and academic achievement (Kashani *et al.*, in press: $X^2 = 2{\cdot}47$, df = 2, $P < 0{\cdot}05$).

Sex of child was not systematically associated with depression, so SPAS Full and subscale scores for these three groups were analysed by means of seven one-way analyses of variance (see Table 4). Post-hoc Scheffé individual comparisons of means were performed where necessary in order to determine where significant effects were occurring. The results show that the depressed group had significantly lower Full Scale scores than either the post-depressed and non-depressed groups. SPAS subscale differences occurred for General Ability, where the depressed group obtained significantly lower scores than either the post-depressed and non-depressed groups, and for Penmanship/Neatness, where depressed group scores were lower than non-depressed group scores.

TABLE 4

SPAS FULL AND SUBSCALE SUMMARY DATA FOR DEPRESSED, PAST DEPRESSED AND NON-DEPRESSED CHILDREN

	Depressed (N = 17)		Past Depressed (N = 23)		Non-Depressed (N = 76)			
	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	F[a]	P
Full Scale	18·73	5·76	23·78	6·37	24·04	5·34	5·71	0·004
General Ability	2·33	2·04	3·04	2·01	3·87	1·70	4·79	0·010
Arithmetic	4·13	2·17	4·74	1·66	4·50	1·70	0·54	0·585
School Satisfaction	3·80	1·52	4·09	1·59	4·17	1·36	0·42	0·655
Reading/Spelling	2·87	1·46	4·30	1·58	4·39	1·79	5·08	0·008
Penmanship/Neatness	4·07	2·05	5·17	1·23	5·22	1·48	3·09	0·028
Confidence	1·53	1·51	2·17	1·67	1·88	1·31	0·95	0·392

[a] df = 2, 111

DISCUSSION

The results of the present study reveal a number of interesting findings about the measurement of academic self-concept as well as the relationships of that construct to other school-related achievement variables, and emotional status.

Sex differences, in which slightly more positive perceptions were reported by girls compared to boys in terms of School Satisfaction, Reading and Spelling, and Penmanship/Neatness were found. These differences were also reflected in the higher mean Full Scale SPAS score obtained by girls. Such a finding has persisted over a number of studies (Boersma and Chapman, 1979; Chapman and Boersma, 1982). Burns (1982) has suggested that the reasons for such differences may arise from the frequent observation that boys appear to have more problems during their primary schooling than girls. The latter tend to achieve better than boys in a wide range of school tasks. Also, Burns (1982) noted that the classroom learning environment tends to be viewed by boys as being " female ", a perception that is reinforced by the predominance of women teachers at the primary level. Consequently, boys may have to cope with more failures (particularly in reading, spelling, and neatness) in a " female " environment, making it more difficult to maintain a positive self-concept.

Turning to the correlational data for academic self-concept with intelligence and school-related performance, some interesting findings emerge. In terms of intelligence, SPAS scores were not meaningfully associated with WISC–R IQ scores. Indeed, apart from some moderate correlations between SPAS General Ability and some WISC–R verbal subtests, Verbal IQ, and Full Scale IQ, all correlations between SPAS and WISC–R scores were weak to negligible. This overall finding is consistent with the very low correlations between SPAS scores and IQ (both individual and group measures) reported in the SPAS Manual (Boersma and Chapman, 1979).

Stronger correlations were noted between the SPAS and the Reading and Spelling tests. This finding is also consistent with other studies where it has been found that SPAS correlations with academic measures tend to be higher for the subject-specific SPAS subscales and corresponding achievement/performance measures than for the other SPAS subscales. Similarly, correlations between parents' ratings of their child's achievement level and SPAS scores also tended to be higher for the subject-specific subscales.

Lesser, moderate to low correlations were also observed between each subject area rated, and both the SPAS Full Scale and General Ability. For the relatively non-subject specific ITPA scales of Comprehension and Expression, however, correlations with SPAS scores were weak to negligible.

In terms of children with specific reading retardation, the results of present study are consistent with numerous other findings where learning disabled children have reported lower perceptions of ability (Chapman and Boersma, 1980; Cullen *et al.*, 1981; Butkowsky, 1982; Hiebert *et al.*, 1982). Although these reading-retarded children may be fairly realistic in their self-perceptions of ability, there is a danger that such relatively negative beliefs will reduce relevant motivation and learning behaviours and hinder efforts aimed at improving reading performance. Indeed, there is a danger that negative self-perceptions of ability will become self-fulfilling prophecies. Thus, persistent reading failure may possibly give rise to self-beliefs of inability to read, low expectations of success on new reading tasks, and possible avoidance of the task or inadequate task-relevant behaviour. As a result, failure is again likely with the consequent confirmation of beliefs about low ability. Where success does occur for children who are failure prone in specific areas, they tend to attribute such success to external sources, such as teacher assistance or task ease, than to their own efforts or ability (Chapman and Boersma, 1979; Pearl *et al.*, 1980).

For the depressed children it is interesting to note that depression is associated with relatively negative self-concepts. In that school represents a major factor in the lives of children, it seems reasonable to expect that depressed children may have misgivings about their abilities in the socially competitive school environment. Alternatively, as Beck (1967, 1974) has postulated, depression may arise from the development of negative self-perceptions, current experience, and foreboding beliefs about the future. Whatever the direction of causal predominance, the association between depression and low self-perceptions of ability would suggest that children with these characteristics are at risk. Developmentally, there is some evidence that self-concept disruption and depressive tendencies most frequently occur during early adolescence (Rosenberg, 1979). For such disruption to be present in 9-year-olds places these children more at risk in resolving issues of identity during adolescence. Failure to do so may result in the development of a diffuse identity and in the inability to satisfactorily establish and maintain mature adult relationships (Erikson, 1965).

Clearly, there are clinical implications of such an interpretation when dealing with depressed children. Active intervention may be warranted in order to deal with the depressive tendencies and to assist in developing more positive self-perceptions of ability.

In summary, this study has demonstrated that a shortened version of the Student's Perception of Ability Scale can be fruitfully used to measure academic self-concept. The results reveal that these self-perceptions were meaningfully related to school achievement, but were relatively independent of measured intelligence. For specific reading retarded children and children diagnosed as depressed, self-perceptions of ability were significantly lower than for normal readers and non-depressed children. School achievement is clearly associated with academic self-concept, and this element should be considered in remedial teaching and in therapy with depressed school children. Overall, Bloom's (1976) belief that academic self-concept is an important factor in school achievement has received further support in this study.

The causal direction in this relationship, however, is still open to question. A logical consideration would suggest that academic self-concept and school achievement are reciprocally interrelated (Bandura, 1977): success in school enhances or maintains academic self-esteem, while self-esteem influences performance through expectations, standards, recognition of personal strengths, higher motivation and level of persistence (Burns, 1982). Conversely, a negative cycle is activated where achievement and/or self-esteem are at a low level. Without doubt, the role of academic self-concept in school achievement appears to be significant.

ACKNOWLEDGMENTS.—The Dunedin Multidisciplinary Health and Development Research Unit is supported by the Medical Research Council of New Zealand, the National Children's Health Research Foundation, the Departments of Education and Health, and involves several departments of the University of Otago, as well as researchers from numerous other universities. Many of the data are gathered by voluntary workers from the Dunedin community. The authors are indebted to the many people whose contributions make this ongoing study possible.

Reprints are available from Dr. James W. Chapman, Education Department, Massey University, Palmerston North, New Zealand.

REFERENCES

ABRAMSON, L. V., SELIGMAN, M. E. P., and TEASDALE, J. (1978). Learned helplessness in humans: critique and reformulation. *J. abnorm. Psychol.*, 87, 49-74.

BANDURA, A. (1977). *Social Learning Theory.* Englewood Cliffs, New Jersey: Prentice-Hall.

BANDURA, A. (1980). Self-referent thought: the development of self-efficacy. In FLAVELL, J. H., and ROSS, L. D. (Eds.), *Cognitive Social Development: Frontiers and Possible Futures.* New York: Cambridge University Press.

BANDURA, A., ADAMS, N. E., HARDY, A. B., and HOWELLS, G. H. (1980). Tests of the generality of self-efficacy theory. *Cognit. Therapy Res.*, 4, 39-66.

BECK, A. T. (1967). *Depression: Clinical, Experimental and Theoretical Aspects.* New York: Hoeber.

BECK, A. T. (1974). The development of depression: a cognitive model. In FRIEDMAN, R. J., and KATZ, M. M. (Eds.), *The Psychology of Depression: Contemporary Theory and Research.* Washington, D.C.: Hemisphere.

BLOOM, B. S. (1976). *Human Characteristics and School Learning.* New York: McGraw-Hill.

BOERSMA, F. J., and CHAPMAN, J. W. (1977). The Student's Perception of Ability Scale. Unpublished instrument. Edmonton: University of Alberta.

BOERSMA, F. J. and CHAPMAN, J. W. (1978). Comparison of the Student's Perception of Ability Scale with the Piers-Harris Children's Self-Concept Scale. *Percept. motor Skills*, 47, 827-832.

BOERSMA, F. J., and CHAPMAN, J. W. (1979). *Manual for the Student's Perception of Ability Scale.* Edmonton: University of Alberta.

BOERSMA, F. J., CHAPMAN, J. W., and MAGUIRE, T. O. (1979). The Student's Perception of Ability Scale: an instrument for measuring academic self-concept in elementary school children. *Educ. psychol. Measur.*, 39, 1035-1041.

BROOKOVER, W. B., LEPERE, J. M., HAMACHEK, D. E., THOMAS, S., and ERICKSON, E. L. (1965). *Self-concept of Ability and School Achievement through Student's Self-concept Enhancement.* U.S. Office of Education, Cooperative Research Project No. 1636. East Lansing: Office of Research and Publications, Michigan State University.

Burns, R. (1982). *Self-concept Development and Education.* London: Holt, Rinehart and Winston.

Butkowsky, I. S. (1982). On the generality of learned helplessness in children with learning difficulties. Unpublished doctoral dissertation, University of Waterloo, Ontario, Canada.

Chapman, J. W., and Boersma, F. J. (1979). Learning disabilities, locus of control, and mother attitudes. *J. educ. Psychol.*, 17, 250-258.

Chapman, J. W., and Boersma, F. J. (1980). *Affective Correlates of Learning Disabilities.* Lisse Holland: Swets and Zeitlinger.

Chapman, J. W., and Boersma, F. J. (1983). A cross-national study of academic self-concept using the Student's Perception of Ability Scale. *NZ J. educ. Studies*, 18, 69-75.

Chapman, J. W., Cullen, J. L., Boersma, F. J., and Maguire, T. O. (1981). Affective variables and school achievement: a study of possible causal influences. *Can. J. behav. Sci.*, 13, 181-192.

Chapman, J. W., Silva, P. A., and Boersma, F. J. (1983). Student's Perception of Ability Scale: development of a Short Form. *Percept. motor Skills*, 57, 799-802.

Covington, M. V., and Beery, R. G. (1976). *Self-worth and School Learning.* New York: Holt, Rinehart and Winston.

Cullen, J. L., Boersma, F. J., and Chapman, J. W. (1981). Characteristics of third-grade learning disabled children. *Learning Disab. Q.*, 4, 224-230.

Erikson, E. (1965). *Childhood and Society.* Harmondsworth: Penguin.

Hiebert, B., Wong, B., and Hunter, M. (1982). Affective influences on learning disabled adolescents. *Learning Disab. Q.*, 5, 334-343.

Hoffman, L. W. (1972). Early childhood experiences and women's achievement motives. *J. soc. Issues*, 28, 129-156.

Kashani, J. H., McGee, R. O., Clarkson, S. E., Anderson, J. C., Walton, L. A., Williams, S. M., Silva, P. A., Robins, A. J., McKnew, D. H., and Cytryn, L. Depression in a sample of nine-year-old children: prevalence and associated characteristics. *Archives of General Psychology*, in press.

Kifer, E. (1975). Relationship between academic achievement and personality characteristics: a quasi-longitudinal study. *Am. educ. Res. J.*, 12, 192-210.

King, L. A., and King, D. W. (1982). Wechsler short forms: a brief status report. *Psychology in the Schools*, 19, 433-438.

Kirk, S. A., McCarthy, J. J., and Kirk, W. D. (1968). *Illinois Test of Psycholinguistic Abilities (Rev. Ed.).* Urbana: University of Illinois Press.

McGee, R. O., and Silva, P. A. (1982). *A Thousand New Zealand Children: Their Health and Development from Birth to Seven.* Dunedin: Medical Research Council of New Zealand. (Special Report Series, No. 8, July, 1982.)

Orvaschel, H., Puig-Antich, J., Chambers, W., Tabrizi, M. A., and Johnston, R. (1982). Retrospective assessment of prepubertal depression with KIDDIE-SADS-E. *J. Am. Academy Child Psychiat.*, 21, 392-397.

Pearl, R., Bryan, T., and Donahue, M. (1980). Learning disabled children's attributions for success and failure. *Learning Disab. Q.*, 3, 3-9.

Purkey, W. (1970). *Self-concept and School Achievement.* Englewood Cliffs, N.J.: Prentice-Hall.

Revicki, D. A. (1982). The relationship between self-concept and achievement: an investigation of reciprocal effects. Paper presented at the annual meetings of the American Educational Research Association, New York, March.

Rosenberg, M. (1979). *Conceiving the Self.* New York: Basic.

Scottish Council for Educational Research (1976). *The Burt Reading Test, 1974 Revision.* London: Hodder and Stoughton.

Shavelson, R. J., Hubner, J. J., and Stanton, J. C. (1976). Self-concept: validation of construct interpretations. *Rev. educ. Res.*, 46, 407-441.

Silva, P. A., Smith, C. T. W., and Pearce, D. W. (1983). A new look at an old spelling test: a report from the Dunedin Multidisciplinary Health and Development Research Unit. Manuscript submitted for publication.

Uguroglu, M. E., and Walberg, H. J. (1979). Motivation and achievement: a quantitative synthesis. *Am. educ. Res. J.*, 16, 375-389.

Wechsler, D. (1974). *The Wechsler Intelligence Scale for Children—Revised.* New York: The Psychological Corporation.

Weiner, B. (1980). The role of affect in rational (attributional) approaches to human motivation. *Educ. Researcher*, 9, 4-11.

Wells, L. E., and Marwell, G. (1976). *Self-esteem: Its Conceptualization and Measurement.* Beverley Hills: Sage.

Wylie, R. C. (1974). *The Self-Concept. Vol. 1: A Review of Methodological Considerations and Instruments.* Lincoln, Neb.: University of Nebraska Press.

Yule, W., Rutter, M., Berger, M., and Thompson, J. (1974). Over- and under-achievement in reading: distribution in the general population. *Br. J. educ. Psychol.*, 44, 1-12.

(*Manuscript received 22nd February*, 1984)

Br. J. educ. Psychol., 54, 293-305, 1984

GETTING A JOB: SCHOOL-LEAVERS' PERCEPTIONS OF EMPLOYMENT PROSPECTS

BY ADRIAN FURNHAM

(*Department of Psychology, University College, London*)

SUMMARY. Whereas most studies of unemployment have tended to focus on the psychological effects of joblessness, few studies have looked at young people's beliefs about, or strategies of, getting a job. This study was concerned with school-leavers' attributions about the causes of unemployment, the most and least useful job-search strategies, and beliefs about which school courses they felt most useful in getting a job. Results showed both sex and class differences in their internal attributions about employment and external attributions about unemployment. Sciences, rather than arts or social science courses, were thought to be most useful in getting a job. The results are discussed in terms of attribution and expectancy theory. Problems associated with research of this kind are also considered.

INTRODUCTION

THE relatively rapid recent increase in unemployment in the western world, and particularly in Great Britain, has led to a great deal of research in this area. A number of topics have been investigated and include the relationship between unemployment and well-being (Hartley, 1980; Warr, 1982), physical and psychological health (Banks *et al.*, 1980; Stafford *et al.*, 1980), and individual difference reactions (Hartley, 1980). There has also been some work on the attributions of both the employed and the unemployed for the causes of unemployment (Furnham, 1982a), the relationship between Protestant ' work ethic ' beliefs and unemployment (Furnham, 1982b) and attitudes toward the unemployed receiving social security benefits (Furnham, 1983).

Most of the above literature has concentrated on adults. However, because of the sudden increase in youth unemployment a number of studies have concentrated on school-leavers. As with adults, most of the studies with school-leavers have looked at the mental health of unemployed young people. Stafford *et al.* (1980) found that young people were at risk of being unemployed if they had an unemployed father, few or no qualifications, low work motivation and if they were black. Furthermore, for young people the psychological consequences may be serious if they do not find a job soon after leaving school. Other studies in the USA (Greenberger *et al.*, 1981) showed similar results. Stafford (1982) looked at the impact of the government sponsored Youth Opportunities Programme on young people's employment prospects and psychological well-being. She found that the programme did act as a buffer, alleviating some of the detrimental psychological effects of unemployment, but that these returned soon after the ex-trainees became unemployed.

As well as a deterioration in mental and physical health there is evidence for various other psycho-social consequences of unemployment on school-leavers. Using a questionnaire developed from the ideas of the neo-psychoanalytic thinker Erikson, Gurney (1980) looked at over 400 unemployed Australian school-leavers. It was hypothesised that having a job helps school-leavers to clarify their perception of their identity and that not being able to get work leads to a confused perception of self or no change due to a moratorium. The hypothesis was confirmed for females but not for males. Furthermore, the unemployed males showed a significant shift towards the mistrust pole of the first dimension of *trust-mistrust* and the employed of both sexes shifted significantly on the *industry-inferiority* subscale. He concludes, ". . . it seems reasonable to conclude that unemployment has the effect of inhibiting development in school-leavers, rather than inflicting trauma as is sometimes popularly supposed " (p. 212). Gurney (1980b) also found that, over a four-month period after leaving school, self-esteem increased only for those young people who obtained work.

More recently, Donovan and Oddy (1982) investigated the social and emotional development of a small, but carefully matched, group of employed and unemployed school-leavers. They found that school-leavers who were unemployed were more depressed and anxious, had lower self-esteem and poorer subjective well-being, were less well socially adjusted and showed a higher incidence of minor psychiatric morbidity than school-leavers who had acquired jobs. They also found a significant interaction between employment status and gender on the locus of control scale—unemployed males tended to be more external than employed males, whereas these differences were minimal for females: They write: " Certainly the apathy and hopelessness frequently associated with unemployment could be linked to an increased tendency to attribute events to uncontrollable forces " (p. 24). Similarly, Feather (1982) found both male and female young, unemployed people had higher depression scores, lower self-esteem and Protestant work ethic beliefs, and reported that good and bad outcomes to everyday events were less important to them compared with employed male subjects. However, Warr *et al.* (1982) found that both psychological distress and self-esteem were uncorrelated with duration of unemployment for young men. They suggested that the expected association between duration and well-being is restricted to older samples and for longer periods out of work.

Considerably less work has gone into establishing young people's beliefs about, and actual attempts at, getting a job. Feather and Davenport (1981) tested their expectancy valence theory on young people: a person's actions are seen to be related to the expectations that a person holds *and* to the subjective values of the outcomes that might occur following the action. They found, as predicted, that higher motivated, more depressed, unemployed youths blamed stable external factors for unemployment and rated the attractiveness of work more highly than less depressed youths. Although incompatible with learned helplessness theory, the results supported their theory which assumes that positive motivation to seek employment is identical to the multiplicative combination of expectancy of success in getting a job *and* the perceived net attractiveness of unemployment. They note, " One might assume that work will have stronger positive valence for individuals with strong Protestant ethic values than for those people for whom these values are weaker. If this assumption is valid, then it follows that people with a strong Protestant ethic value will be more persistent in their efforts to get a job and that they will suffer more negative effects if they fail to obtain employment " (p. 337).

Gurney (1981) in fact examined the attributions for the causes of unemployment in both employed and unemployed groups of school-leavers. In a first study he attempted to discover, among a population of Australian school-leavers, whether the unemployed differed in their attributions of employment from those who succeeded in getting work and, secondly, whether any differences were antecedent to, or consequent upon, unemployment. Subjects were asked to ascribe the ability to get work either to internal or external factors to the job seekers, and to complete a simple 8-item scale devised by the author. He found that unemployed males attributed both getting *and* not getting work significantly more to external factors than did the employed males, but that there were no differences between the female group. Gurney suggested that the lack of differences between the female groups may be due to their lesser defensiveness and need to blame external factors for being unemployed.

In a second longitudinal study students were given a 12-item attribution-for-getting-jobs questionnaire prior to leaving school and then approximately four months later. He suggested that: " The unemployed may believe themselves powerless to change the circumstances of their lives (external locus) but the fact that they are without work whereas others are not, and that they remain so, may lead them eventually to see themselves as responsible for their condition (internal attribution) with consequent self-blame and self-derogation " (p. 89). The results showed that prior to leaving school the groups did not differ, yet later it was not the unemployed who changed their attributions, but the employed who shifted significantly toward a more internal set of causal ascriptions. Overall Gurney (1981) seemed unable to account for his " counter-intuitive and unexpected " findings, suggesting perhaps that subjects of this age have their self-esteem based on numerous other things such as parental evaluation and peer group approval, rather than exclusively on work which is more often the case with adults.

In a rather different study Dayton (1981) looked at the way in which young people looked for a job. He set out to determine what job-seeking approaches were being used by young people and what factors they found positive and useful (Aids) and what negative and worthless (Barriers) in a job search. Using a population of 250 young Americans, Dayton found they regarded their own positive personal attributes (personality, flexibility, academic ability) as the most important aids to their job search, and external factors (labour unions, welfare and unemployment insurance, government training programmes) as least important. Employment success and satisfaction were correlated with careful analysis of which job suited them best, the assemblance of a placement file, letters of recommendation, and a curriculum vitae, combined with persistence in the job search.

Research, within the framework of attribution theory, would, however, lead one to make a number of predictions about schoolchildren's expectations and beliefs about getting a job (Furnham, 1982a, 1982b, 1983). For instance, people more prone to unemployment, and the unemployed themselves, tend to make more external attributions for the causes of unemployment in contrast to those in jobs and unlikely to become unemployed. Furthermore, studies have shown that external attributions are to some extent protective of self-esteem in the context of achievement (Harvey *et al.*, 1978; Weiner *et al.*, 1978). Hence it may be predicted that females more than males and working class subjects more than middle class subjects—for whom unemployment is statistically more probable—will be prone to make more external attributions about getting a job. These attributions will also be reflected in the number and type of job-search strategies adopted by young people and the barriers and aids that they consider operate in job-hunting success.

This study set out to examine sex and class differences in school-leavers' attributions about unemployment, the most and least useful job-search strategies, and which school courses they believed most and least useful in getting a job.

METHOD

Subjects

Two groups of subjects were tested. The first comprised 134 subjects (49 male, 85 females) from a middle class school in the London commuter belt. Their median age was 16 years and nearly all were sitting their GCE A-levels in a variety of subjects. The second group comprised 103 subjects (54 males, 49 females) from a working class school in the East End of London (median age 16). An analysis of the school's catchment areas and teachers' reports confirmed the class categorisation. As well as age and sex, information was gathered on which A-levels each candidate was doing, whether their parents were employed or not and what they would vote if given the opportunity. An analysis of the parental occupations (where provided in sufficient detail) provided evidence of the class classification: middle class children came from class I, II, and IIIa homes while the working class children came from class IIIb and IV homes.

Questionnaire

Each subject completed two questionnaires each of which had two parts:

(1) *Attribution Questionnaire* (Gurney, 1981)
This was a two-part questionnaire developed to determine to what extent school-leavers believe the ability/inability of school leavers to get work is mainly a function of internal or external factors. The first questionnaire had 8 items and the second, which was a development of the first, had 12. Whereas the first had a modest α coefficient of 0·442 (N = 131) the latter had an α coefficient of 0·54 (N = 434). Although Gurney found some counter-intuitive results the scales appear to be both reliable and valid.

(2) *Job-seeking Activities* (Dayton, 1981)
This was a two-part questionnaire developed to determine firstly, which of 30 or so activities were considered to be most useful getting a job and secondly, which of 35 factors were considered barriers and which aids to getting a job. Although Dayton did not do much psychometric work on these scales (validity, reliability) he did administer them to two populations with comparable results and found they correlated meaningfully with employment status and satisfaction and a wide range of demographic and aptitude measures.

(3) *The most and least useful A-levels for actually " getting a job "*
Subjects were also requested to list (up to 3 in each case) what they considered to be the most and least useful A-levels for getting a job. Following previous categories the following were considered *science* subjects: biology, chemistry, physics, geology, geography, mathematics, computer science, psychology, zoology, social biology and statistics. The following were considered *art* subjects: English, Latin, Greek, modern languages, history, art, music, religious studies, and the following *social sciences*: sociology, politics, economics, economic history.

Procedure

Subjects were tested in the class time by their teachers. The questionnaires took approximately 25 minutes and subjects were fully debriefed at a later stage. All appeared to understand the instructions and fewer than 5 per cent of the questionnaires were spoiled or incomplete.

RESULTS

Five independent variables were examined in this study: sex, class, A-level, voting intentions, and parents' employment status. Thus five sets of ANOVAs were computed across all the items on both the Gurney (1981) and Dayton (1981) scale. The only variables which discriminated better than chance were sex and class. There were, however, class differences in the subjects' beliefs about sex and racial advantage in getting a job. The working class sample believed males are more likely than females ($F = 10{\cdot}04$, $P < 0{\cdot}001$) and whites more likely than blacks ($F = 9{\cdot}98$, $P < 0{\cdot}001$) to get jobs.

(A) *Attributions for employment*

The results of the 2-way ANOVA (sex × class) for the 20 attribution items is shown in Table 1.

From Table 1 it can be seen that among the items that subjects most *disagree* with are (3): " Most kids get jobs if they are confident, look hard and have a lot to offer ";

TABLE 1

MEANS AND F LEVELS FOR THE 2-WAY (SEX × CLASS) ANOVA ON THE 20 ATTRIBUTION ITEMS

	Male		Female		F Levels		
	Working Class (N=54)	Middle Class (N=47)	Working Class (N=60)	Middle Class (N=85)	Sex	Class	Sex × Class
1. It is mainly a matter of luck whether a school-leaver gets a job or not.	3·11	3·23	3·42	3·18	0·65	0·14	1·21
2. If they had better qualifications most of the unemployed would soon get jobs.	2·00	2·48	1·92	2·22	1·31	6·86**	0·38
3. Most kids get jobs if they look hard, are confident and have a lot to offer.	3·89	3·51	3·84	3·83	0·87	1·89	1·80
4. Most school-leavers who have got a job had someone " pulling strings " for them.	3·11	2·76	2·42	2·61	7·14**	0·25	2·91
5. Getting a job is mainly a matter of being in the right place at the right time.	2·23	2·44	2·38	2·23	0·04	0·84	1·19
6. Unemployment is running so high because the jobs simply are not there.	3·03	2·65	2·82	2·78	0·07	1·65	1·18
7. It is mainly the " smarter " kids who have been able to get work.	3·55	3·34	3·18	3·40	0·92	0·00	1·73
8. Unemployed kids haven't tried hard enough and don't know how to sell themselves.	3·91	3·34	3·89	3·44	2·27	5·91**	1·16
9. Young people miss out on getting jobs because employers are prejudiced against them.	3·52	3·04	3·24	3·36	1·27	0·10	1·55
10. If you miss out on getting work it is because you are not good enough.	2·42	2·78	2·46	2·68	0·07	5·35*	0·31
11. Kids get work if they look hard & often enough.	3·17	2·55	2·94	2·68	0·12	7·83**	1·26
12. The government is to blame for young people being out of work.	3·66	3·70	3·34	3·49	3·59*	0·48	0·69
13. Getting a job depends on sheer good luck.	3·86	3·63	3·40	3·31	9·56***	1·48	0·32
14. School-leavers are unemployed because older people have taken all the jobs.	2·62	2·72	2·54	2·70	0·17	1·08	0·08
15. Young people don't get jobs because they are not good at " putting themselves over " in applications and interviews.	2·62	2·63	2·98	2·75	3·17	0·68	0·82
16. You can get a job if you are well qualified.	3·81	3·68	3·76	3·87	0·29	0·01	0·94
17. People who haven't got work don't really want to work or haven't looked hard enough.	3·50	3·31	3·64	3·21	0·01	5·37*	0·80
18. If you're good looking and have lots of confidence you'll get a job.	3·13	3·23	3·24	3·23	0·18	0·14	0·17
19. To get a job, you need someone with influence to " put in a good word " for you.	2·45	2·72	2·46	2·52	0·47	1·43	0·44
20. Going to the right school and having the right contacts is a big part of getting a job.	2·71	2·99	2·41	2·87	1·19	5·14*	0·86

* P<0·05 ** P<0·01 *** P<0·001

The numbers represent the mean on the scale 1 = always true to 5 = never true.

(8): " Unemployed kids haven't tried hard enough and don't know how to sell themselves "; and (16): " You can get a job if you are well qualified ". On the other hand, among the items most subjects *agreed* with are (2): " If they had better qualifications most of the unemployed would soon get jobs "; (5) " Getting a job is mainly a matter of being in the right place at the right time ". On the whole the subjects tended to believe that effort, perseverance, and help from others are most important in getting work and luck least important. However, they also believe that people do not have jobs because they are simply not there, and that unemployed people want to work and look for work. Interestingly, subjects tended to agree with internal or personal attributions for getting a job, but when given similar reasons for the inability to get jobs they tended not to agree. This appears to imply that the factors for getting a job are not necessarily the same as those which lead one not to get a job (Gurney, 1981).

There were three significant sex differences. Females more than males tended to believe in external factors: someone pulling strings for them, government policy, and sheer good luck. There were six significant class differences. Working class, more than middle class subjects, believed that if they had better qualifications most of the unemployed would soon get jobs; if you miss out on getting work it is because you are not good enough and that going to the right school and having the right contacts is a big part of getting a job. On the other hand the middle class sample, more than the working class sample, believed that unemployed school-leavers have not tried hard enough and do not know how to sell themselves; that they get work if they look hard and often enough; and that people who haven't got work don't really want to work or haven't looked hard enough. Therefore, whereas the working class tend to stress external and educational factors, the middle class tend to stress internal and effort factors.

(B) *Strategies*

The results of the sex × class ANOVA on the 30 strategies are shown in Table 2. Most of the 30 items were considered useful. In fact, only three had a mean score of above 3·00: (14) Seek help from your family (3·14); (15) selecting a specific job to look for (3·77); and (17) Move to a better employment area (3·07). Knocking at potential employers' doors, taking a civil service exam, and taking a vocational training scheme also ranked in the bottom six. Strategies considered among the most useful included taking a job interest/aptitude test, writing letters of inquiry, assembling a placement file, talking to a school/employment counsellor, visiting companies and employers, doing summer or after-school work, and finding out what courses are most helpful for getting jobs.

Six sex differences reached significance: males considered assembling a placement file, seeking help from their family and doing volunteer work more useful than did females who considered applying to a governmental agency, visiting companies and employers and doing summer or after-school work. Overall, these results suggest that females place more emphasis on gaining experience and males on applications *per se*.

All of the eight class differences showed that the middle class found certain strategies more useful than the working class: telephoning potential employers, registering with a private employment service, assembling a placement file, selecting a specific job and looking for it, studying information on job trends, placing an advertisement in the newspaper, taking job interview training and taking vocational training. There were three significant interactions: working class males and middle class females more than working class females or middle class males believed seeking help from teachers and friends, gathering and using letters of recommendation, and studying job trends were useful strategies for getting a job.

TABLE 2

MEANS AND F LEVELS FOR THE 2-WAY (SEX × CLASS) ANOVA ON THE 30 APPLICATION ITEMS

	Male		Female		F Levels		
	Working Class	Middle Class	Working Class	Middle Class	Sex	Class	Sex × Class
1. Writing up & sending out your CV.	2·47	2·48	2·58	2·16	0·80	2·68	3·10
2. Seeking help from friends.	2·18	1·97	2·02	1·95	0·71	1·45	0·38
3. Search the classified want adverts.	2·52	2·48	2·76	2·67	2·32	0·21	0·04
4. Take a job interest/aptitude test.	1·67	1·61	1·74	1·62	0·10	0·68	0·07
5. Write letters of inquiry.	1·82	1·85	2·16	1·91	1·67	1·30	0·59
6. Telephone potential employers.	2·08	1·80	2·20	1·92	0·86	4·61*	0·00
7. Seek help from teachers or lecturers.	2·69	2·27	2·24	2·35	2·02	1·31	3·97*
8. Register with the state employment service.	2·55	2·36	2·32	2·41	0·49	0·15	1·15
9. Register with a private employment service.	2·91	2·40	2·70	2·35	1·18	12·23***	0·45
10. Register with a school/college placement service.	2·64	2·63	2·64	2·29	1·70	1·73	1·62
11. Assemble a placement file of jobs going.	1·88	1·63	2·18	1·88	4·72*	4·69*	0·05
12. Gather & use letters of recommendation	3·03	2·44	2·98	3·04	2·51	2·28	3·60*
13. Knock at potential employers' doors.	2·57	2·44	2·78	2·56	1·36	1·56	0·10
14. Seek help from your family.	2·94	2·74	3·56	3·35	18·13***	2·07	0·00
15. Select a specific job and look for it.	3·91	3·59	4·00	3·63	0·19	5·84**	0·03
16. Join or try to join a trade union.	2·61	2·57	2·92	2·72	2·69	0·64	0·30
17. Move to a better employment area.	3·01	2·97	3·12	3·18	1·02	0·00	0·05
18. Take a civil service exam.	3·11	2·95	2·76	1·92	2·60	0·00	1·90
19. Apply to some government agency.	2·27	2·31	2·10	1·89	4·77*	0·33	0·86
20. Do volunteer work related to chosen job.	2·55	2·36	2·84	2·62	3·69*	2·15	0·00
21. Study information on job trends.	3·35	2·61	2·88	2·78	0·91	7·24**	4·34*
22. Place an advertisement in the newspaper.	2·59	2·04	2·48	2·34	0·56	7·75**	2·76
23. Analyse your interests & abilities.	2·04	2·10	2·04	2·05	0·06	0·09	0·02
24. Take job interview training.	2·69	2·08	2·66	2·14	0·01	18·30***	0·12
25. Take vocational training.	3·08	2·57	2·84	2·58	0·66	7·95**	0·84
26. Seek help from groups that you belong to	2·20	2·25	2·24	2·30	0·17	0·22	0·00
27. Talk to a school/employment counsellor.	1·71	1·95	1·90	1·87	0·21	0·98	1·58
28. Visit companies & employers to learn more.	1·76	1·95	1·50	1·65	7·61**	3·02	0·03
29. Do summer or after-school work for experience.	1·87	1·85	1·58	1·75	4·10*	0·37	1·15
30. Find out what courses are most helpful for jobs.	1·77	1·98	1·76	1·81	0·14	0·81	0·41

$* P < 0·05$ $** P < 0·01$ $*** P < 0·001$

The numbers represent the mean on the scale 1 = extremely useful to 7 = not at all useful.

(C) *Aids and Barriers*

The results of the sex × class ANOVA on the 35 aids and barriers are shown in Table 3. Overall the five most useful aids were considered to be (20) your persistence/determination in looking for work (1·59); (14) your work qualifications (1·64); (5) your personality (1·66); (2) your reading and writing ability (1·66); and (17) your ability to get along with bosses and co-workers (1·69). The least useful were considered to be item (26) your lack of desire to find work (4·44); (33) the unemployment rate (4·25); not knowing what you want to do (4·24); (31) any physical disability you may have (4·09); (18) your previous problems with the law (4·09). Thus all five of

TABLE 3

Means and F Levels for the 2-way ANOVA (Sex × Class) on the 35 Barriers/Aids to Getting a Job

	Male		Female		F Levels		
	Working Class	Middle Class	Working Class	Middle Class	Sex	Class	Sex × Class
1. The state of the British economy.	4·02	3·82	4·00	3·94	0·13	0·76	0·23
2. Your reading and writing ability.	1·85	1·75	1·56	1·55	3·90*	0·18	0·12
3. Counselling/job placement help that you received.	2·70	2·34	2·31	2·49	1·27	0·65	6·19*
4. Trade unions.	3·14	3·63	3·37	3·64	0·79	7·77***	0·61
5. Your personality.	1·72	1·85	1·41	1·03	6·17*	2·77	0·16
6. Your not knowing what you wanted to do.	4·20	4·24	4·31	4·21	0·09	0·06	0·33
7. Your race.	3·35	3·57	3·31	3·38	0·55	0·97	0·23
8. Employers' attitudes toward young people.	3·40	3·67	3·54	3·35	0·67	0·11	3·97
9. The kinds of jobs available.	3·11	3·18	3·20	3·12	0·03	0·00	0·56
10. The availability of unemployment insurance.	3·05	3·20	3·00	3·15	0·24	1·89	0·00
11. Your sex.	2·87	2·87	3·45	3·43	18·83***	0·00	0·01
12. Your high job expectations.	2·77	2·95	3·06	3·09	1·90	0·49	0·24
13. Vocational training that you have received	2·35	1·97	1·95	1·82	6·46*	5·50*	1·21
14. Your work qualifications.	1·68	1·71	1·56	1·62	1·00	0·18	0·02
15. Money given to the Social Security.	3·11	2·91	2·89	2·98	0·55	0·26	2·12
16. Your maths ability.	1·74	2·08	1·58	2·09	0·34	11·59***	0·46
17. Your ability to get along with bosses/co-workers	1·72	1·70	1·56	1·75	0·24	0·62	1·13
18. Your previous problems with the law.	4·17	3·71	4·67	3·94	5·18*	13·60***	0·73
19. Your general health.	2·31	2·51	2·62	2·20	0·00	0·56	4·10*
20. Your persistence/determination in looking for work	1·54	1·71	1·52	1·59	0·50	1·47	0·30
21. Help from friends and personal contacts.	1·93	1·96	1·98	1·96	0·07	0·01	0·05
22. Your willingness to work part time/odd hours.	1·85	1·86	1·79	1·72	0·95	0·11	0·15
23. Your experience in organisations (Scouts, etc.)	2·44	2·26	2·42	2·12	0·57	4·23*	0·27
24. Your high school record.	1·80	1·94	1·58	1·80	2·22	2·02	0·07
25. Competition for the job that you want.	3·80	3·98	4·02	3·92	0·35	0·08	1·09
26. Your lack of desire to find a job.	4·67	4·10	4·67	4·35	1·47	18·06***	1·47
27. Your age.	3·33	3·33	3·58	3·30	0·00	3·50	3·32
28. Your parents' background and help.	2·57	2·41	2·65	2·30	0·03	3·89*	0·50
29. The city & neighbourhood in which you live.	1·13	2·59	3·40	2·67	1·99	18·06***	0·73
30. Government sponsored employment programmes.	3·06	3·02	2·69	2·75	6·07*	0·01	0·15
31. Any physical disability that you have.	4·09	3·82	4·38	4·09	5·35*	5·29*	0·00
32. Attitudes of your friends about work.	3·26	2·84	3·00	3·05	0·05	2·84	4·45*
33. The unemployment rate.	4·30	4·22	4·08	4·32	0·27	0·49	1·75
34. Your knowledge of the kinds of work possible.	2·55	2·37	2·52	2·15	0·93	4·65*	0·49
35. Chance and luck.	2·44	2·46	2·25	2·33	1·62	0·16	0·04

* $P<0{\cdot}05$ ** $P<0{\cdot}01$ *** $P<0{\cdot}001$

The numbers represent the mean on the scale 0 = very useful to 5 = not very useful.

the most useful items were internal individualistic factors, while most, but not all, of the least useful items were external or societal.

Seven of the items showed sex differences. Reading and writing ability and personality were seen as greater aids by females than by males as were vocational training and government sponsored employment programmes. Females too saw physical disability as less of a barrier than did males but males saw problems with the law as being more of a hindrance. The largest difference, however, was an item referring to sex differences: males saw this as more of a help than a hindrance, whereas to females it was the reverse. Thus both groups were aware of some form of sex discrimination in the seeking of employment.

There were 10 significant class differences. The working class saw vocational training and experience in organisations as being significantly less helpful than did their middle class counterparts. Parents' background and help were considered less of an aid; the neighbourhood in which they lived was considered as more of a hindrance by the working class than the middle class sample. The working class children also found knowledge of the work available as less helpful. Physical disability and, particularly, problems with the law were seen as greater handicaps by the working class group. The only factor that the middle class saw as more of a barrier was that of trade unions.

There were three significant interactions. Whereas working class females found counselling/job-placement help more of an aid than either of the middle class groups, working class males found it least helpful. On the other hand, ' general health ' was considered to be far more of an asset by middle class females and working class males than the other two groups. Finally, whereas middle class males found the attitudes of friends towards work neither a help nor a hindrance, the other groups found this item a barrier with working class females finding it most unhelpful. In summary it appears that girls and the working class sample saw certain items as less helpful than boys and the middle class groups.

(D) *Most and least useful A-levels*

Table 4 gives the A-levels most nominated by the subjects as being most and least useful. All of the top five ranked items save two (English and economics) were

TABLE 4

THE MOST AND LEAST USEFUL A-LEVEL SUBJECTS FOR GETTING A JOB

A. Most Useful		
	Number of nominations	% of all nominations
1. Mathematics	209	32·15
2. English	152	23·38
3. Physics	95	14·61
4. Chemistry	32	4·92
5. Economics	30	4·61
B. Least Useful		
	Number of nominations	% of all nominations
1. History	70	12·41
2. Art	61	10·81
3. Classics	42	7·46
4. Religious Education	42	7·46
5. Geography	39	6·91
6. Sociology	39	6·91
7. Cookery	38	6·73

C. Sex Differences in nominations of most useful		
	Male	Female
Arts	67	133
Science	191	211
Social Science	25	23

science courses. Mathematics was by far the most popular choice in that nearly a third of all subjects chose this course. There was less agreement about what constituted the least useful A-level courses. History was considered to be the least useful followed by art, classics, and religious education. Geography and sociology were also rated as fairly useless in getting jobs. Overall, science courses formed 62 per cent of the subjects considered most useful. Arts courses were mentioned 31 per cent of the time and social science/miscellaneous just 7 per cent. However, when looking at the least useful nominations 13 per cent mentioned science courses, 64 per cent arts courses, and 23 per cent social science courses.

The total responses for the ' most useful ' A-levels were tabulated for sex and class; a χ^2 was performed. There were no class differences but a significant sex difference ($\chi^2 = 10{\cdot}96$, $P < 0{\cdot}05$). Females showed a greater tendency to select arts courses as helpful for jobs than did males. Males selected arts courses only 25 per cent of the time, whereas females selected arts courses 37 per cent of the time. There were no sex or class effects in those subjects nominated as least useful.

It was also possible to check the relationship between the A-levels actually being undertaken by the subjects and the course they themselves were following. Nearly 35 per cent of the males, as opposed to 14 per cent of the females, followed science-only courses whereas 21·5 per cent of the females, as opposed to 7·5 per cent of the males, followed arts-only courses. Thus it is that subjects tend to believe that the course that they are doing is the most useful in getting a job. Of course the direction of causality cannot be ascertained from this data—do females follow more arts-based courses because they do not perceive them to be as useless as do males or is it more the case that having selected these courses they are more prone to want to believe they are useful?

DISCUSSION

The results of the four different parts of this study suggest that, overall, attributions about getting a job are frequently internal (that is to personal attributes or abilities rather than environmental or societal factors). Confidence, perseverance, and qualifications were all considered to be primary factors responsible for success in finding employment, yet this was moderated by the belief that jobs are not currently available (a fact which is attributed to the government). Yet failure to get a job was rarely attributed to the personal shortcomings of the job-seeker himself. Thus these results tend to support the well-established, attributional finding that success is attributed to internal factors and failure to external factors. These school-leavers did not show the protective external attributions found among the unemployed when considering success in finding employment (Furnham, 1982a, 1982b) but seem to be covering themselves for the possibility of joblessness. This may be seen as an example of defensive or self-serving attribution.

Where there were significant sex differences it was found that females were more external in their attributions than males. This conflicts with Gurney (1981) who found that unemployed males were more external in their attributions about employment than employed males, but that there were no differences between employed and unemployed females. However, the extensive locus of control literature has shown that, where sex differences exist in generalised locus of control beliefs, females are more external than males. Similarly, class differences tended to show that working class subjects tend to place more emphasis on structural or external factors (Furnham, 1982a). Again this is to be expected and in accordance with previous literature on the topic (Furnham and Gunter, 1984).

The results on the aids and barriers confirms the findings of Dayton (1981) who found that the subjects saw their own personal attributes as the greatest aids and external factors as the biggest hindrances. Similarly, regarding strategies, class and sex differences showed that the middle class tended to rate all job-hunting strategies as more useful than the working class and girls showed less faith in following up specific job choices than did boys. The subjects all stressed the importance of summer and after-school work for experience but tended to rely on personal contacts rather than direct approaches to employers. It would be interesting to compare these beliefs with those of employers who may have quite different beliefs concerning which factors make an applicant more employable.

The belief in the usefulness of A-level courses revealed that both males and females believed that science courses (and English) were the most useful in getting a job, although females tended to opt for arts courses and males for science courses at A-levels. Females also believed that arts courses were more useful than science courses, so providing a rationalisation for their choice. However, a study such as this was unable to determine whether females chose arts subjects and then felt it necessary to justify their choice or whether they actually believed them to be most useful *per se* and hence selected them accordingly. It is, of course, also possible that when candidates select a particular course they do so for many reasons only one of which is its usefulness for getting a job. Further, the believed importance of science courses may reflect recent government emphasis and finding towards the " hard " sciences rather than the arts or social sciences.

These results also indicate that attributions are to some extent a function of the expectation of work. Furnham (1982a, 1982b) and Gurney (1981) have demonstrated that the unemployed are more external in their employment attributions and it also appears that those who are statistically more likely to be unemployed (females *vs.* males, and working class *vs.* middle class school-leavers) are more external in their attributions than those who expect to get a job. These results are clearly interpretable in terms of expectancy theory which suggests that the strength of a person's motivation to seek work is related to his or her expectations that job-seeking efforts will result in employment *and* to the perceived attractiveness of the nature of the employment itself. " The expectations may be determined by a variety of different factors that include as especially important ones a person's perceptions of his or her ability and skill and of the ease or difficulty of the present task. People who see themselves as relatively low in competence, for example, and who perceive that few jobs are available would tend to have lower expectations of success " (Feather and Davenport, 1981, p. 423).

However, it is not clear whether lower expectations of success act as a self-fulfilling prophecy in that those who perceive themselves (rightly or wrongly) as being low in ability and qualifications do not adopt job-search strategies that would actually help in getting a job. On the other hand, lower expectations might accurately reflect the job-market at the time. Nevertheless, the results of this study do suggest that job-search strategies are related to attributions about unemployment and that some, at least, of the lack of success of school-leavers to get jobs may be explicable in terms of their inappropriate strategies. However, despite or perhaps because of low expectations of a job there appears to be no decline in young people's willingness to work. Borus (1982) found in a large group of Americans (aged 14-22) that many were very willing to work at sub-minimum wages indicating their desire to gain work experience.

An adolescent's beliefs about aids and barriers to finding a job may also relate to their behaviour in job interviews. Recently there has been a great deal of work on job interview training with adolescents in Britain (Hood *et al.*, 1982) and America (Heimberg *et al.*, 1982). As well as giving training in non-verbal (eye contact, gesture) and verbal (questioning, answering) behaviours the therapists strove to change the attitudes and beliefs of the applicants. Nearly all studies have shown that social skills training for job interviews improves both role play performance and the probability of getting jobs. This aid was considered by Dayton (1981) and may be worth including in future studies in this area.

This study was not without its limitations. The samples were not particularly large or well-matched, and the class classification was fairly crude. Also, whilst the samples tested cannot be taken as representative of the British population of school leavers, there is no reason to believe that they were atypical. They certainly provide useful comparisons for present purposes although it may be unwise to generalise these results to all school-leavers. Furthermore, this study concentrated on the hypothetical aspects of the subjects' job-search strategies rather than on their actual strategies. Nevertheless, the hypotheses that were entertained received some empirical support and confirmed the results of previous studies in this important and neglected area of research.

Acknowledgment.—Extensive help was received from Caroline Muller in the analysis of the data.

Reprint requests should be directed to Dr. Adrian Furnham, Department of Psychology, University College, London, 26 Bedford Way, London, WC1E 6BT.

REFERENCES

Banks, M., Clegg, C., Jackson, P., Kemp, N., Stafford, E., and Wall, T. (1980). The use of the General Health Questionnaire as an indicator of mental health in occupational studies. *J. occup. Psychol.*, 53, 187-194.

Borus, M. (1982). Willingness to work among youth. *J. hum. Resources*, 17, 581-593.

Dayton, C. (1981). The young person's job search: insights from a study. *J. Counsell. Psychol.*, 28, 321-333.

Donovan, A., and Oddy, M. (1982). Psychological aspects of unemployment: an investigation into the emotional and social adjustment of school leavers. *J. Adolescence*, 5, 15-30.

Feather, N. (1982). Unemployment and its psychological correlates: a study of depressive symptoms, self-esteem, protestant ethic values, attributional style and apathy. *Aust. J. Psychol.*, 34, 309-323.

Feather, N., and Davenport, P. (1981). Unemployment and depressive effect: a motivational and attributional analysis. *J. person. soc. Psychol.*, 41, 422-436.

Furnham, A. (1982a). Explanations for unemployment in Britain. *Eur. J. soc. Psychol.*, 12, 335-352.

Furnham, A. (1982b). The Protestant work ethic and attitudes towards unemployment. *J. occup. Psychol.*, 55, 277-286.

Furnham, A. (1983). Attitudes to the unemployed receiving social security. *Human Relations*, 36, 135-150.

Furnham, A., and Gunter, B. (1984). Just world beliefs and attitudes towards the poor. *Br. J. Soc. Psychol.*, 23, 265-269.

Greenberger, E., Steinberg, L., and Vaux, A. (1981). Adolescents who work: health and behavioural consequences of job stress. *Dev. Psychol.*, 17, 691-703.

Gurney, R. (1980a). The effects of unemployment on the psychosocial development of school leavers. *J. occup. Psychol.*, 53, 205-213.

Gurney, R. (1980b). Does unemployment affect the self-esteem of school leavers? *Aust. J. Psychol.*, 32, 175-182.

Gurney, R. (1981). Leaving school, facing unemployment and making attributions about the causes of unemployment. *J. vocat. Behav.*, 18, 79-91.

Hartley, J. (1980). The personality of unemployed managers: myths and measurements. *Personnel Rev.*, 9, 11-18.

Harvey, J., Ickes, W., and Kidd, R. (Eds.) (1978). *New Directions in Attribution Research* (Vol. 2). Hillsdale, N.J.: Erlbaum Press.

Hayes, J., and Nutman, P. (1981). *Understanding the Unemployed.* London: Tavistock.

Heimberg, R., Cunningham, J., Stanley, J., and Blankenberg, R. (1982). Social skills training to prepare unemployed youth for the job interview: a controlled evaluation. *Behav. Modification*, 6, 299-322.

Hood, E., Lindsay, W., and Brooks, N. (1982). Interview training with adolescents—a controlled group study incorporating generalization and social validation. *Behav. Res. Therapy*, 20, 581-592.

KELVIN, P. (1980). Social psychology 2001: social psychological bases and implications of structural unemployment. In GILMOUR, R., and DUCK, S. (Eds.), *The Development of Social Psychology.* London: Academic Press.

STAFFORD, E., JACKSON, P., and BANKS, M. (1980). Employment, work involvement and mental health in less qualified young people. *J. occup. Psychol.*, 53, 291-304.

STAFFORD, E. (1982). The impact of the youth opportunities programme on young people's employment prospects and psychological well-being. *Br. J. Guidance Counsell.*, 10, 12-21.

WARR, P. P. (1982). Psychological aspects of employment and unemployment. *Psychol. Med.*, 12, 7-11.

WARR, P., JACKSON, P., and BANKS, M. (1982). Duration of unemployment and psychological well-being in young men and women. *Current Psychol. Res.*, 3, 207-214.

WEINER, B., RUSSELL, D., and LERMAN, D. (1978). Affective consequences of causal ascriptions. In HARVEY, J., ICKES, W., and KIDD, R. (Eds.), *New Directions in Attribution Research* (Vol. 2). Hillsdale, N.J.: Erlbaum.

(*Manuscript received*, 6*th February*, 1984)

Br. J. educ. Psychol., 54, 306-319, 1984

IDENTIFYING DIMENSIONS OF SCHOOL MOTIVATION IN BRITAIN AND HUNGARY

BY B. KOZÉKI
(Institute of Psychology, Budapest)
AND N. J. ENTWISTLE
(Department of Education, University of Edinburgh)

SUMMARY. Previous research in Britain and USA has indicated that motivation in education has several dimensions. Besides academic achievement motivation (hope for success), there are also dimensions reflecting fear of failure, extrinsic motivation related to external rewards, and intrinsic motivation out of personal interest in the subject matter itself.

In Hungary extensive interview studies have led to the development of an integrated description of school motivation involving nine distinguishable motives in affective, cognitive, and moral domains. An inventory was developed to measure these nine hypothesised motives within the school context, and six separable factors were identified in Hungary.

The present study involves a comparison of the factor structure of the Hungarian inventory with an English-language version. Eight hundred Hungarian and 365 British 12-13-year-olds completed the inventory, and the main analysis extracted eight factors for boys and girls in each country. These factors described dimensions similar to the hypothesised motives, which were identifiable in a comparable form in both samples. Subsidiary analyses examined sex and school differences in mean scores within the two countries. Examination of individual items also produced interesting indications of possible differences between the reactions of British and Hungarian children to their experiences of school.

INTRODUCTION

AMONG the many attempts to describe motivation in relation to educational processes, a fundamental distinction can be made between studies which utilise a general theoretical framework and those which try to build theory more directly out of classroom events and experiences.

The first approach is perhaps best represented by the work of Cattell (1957) and his co-workers, who sought to explain behaviour in a wide range of social settings in terms of fundamental psychological drives (ergs) and sentiments. Although this approach has produced interesting findings in relation to school attainment (Cattell and Child, 1975), the indicators of motivation are mainly remote from the specific context of the classroom.

The other main approach to school motivation, which is followed here, has involved the development of inventories which use items relating to classroom experiences or which are specifically designed to predict academic attainment. Early inventories, both in the USA (Finger and Schlesser, 1965) and in Britain (Entwistle, 1968), concentrated on the development of unidimensional scales of *academic (achievement) motivation*. These early scales did indeed explain additional variance in attainment beyond that related to verbal reasoning scores, but their conceptual bases were far from clear. In fact, their role in the prediction of attainment was disparagingly likened to that of a " conceptual charlady " mopping up unexplained variance (Jones *et al.*, 1973). Subsequently, it was recognised that conceptual clarification was essential for several different types of motivation related to education (Entwistle *et al.*, 1974). There were, at least, extrinsic motivation (related to external rewards) and two forms of intrinsic motivation—one related to a personal interest in the subject matter, and the other to self-esteem. The latter form of motivation could be identified with ' n-ach ' (McClelland *et al.*, 1953) or ' hope for success ' (Atkinson and Feather, 1966), and distinguished from ' fear of failure ' (Birney *et al.*, 1969). In work on the academic performance of students, three distinct types of successful students were identified (Entwistle and Wilson, 1977). Each group of students seemed to follow a different path to success which was dependent on contrasting motivational styles (Kozéki and Entwistle, 1983). The first group appeared to be motivated by ' hope for success ', the second by ' fear of failure ', and the last by intrinsic motivation rooted in personal interest and syllabus-freedom. In very recent work on students these three types of motivation, together with extrinsic (vocational) motivation, were again identified, and were found to be related to contrasting approaches to studying (Entwistle and Ramsden, 1983).

In the USA, Ball (1977) has described many of the different theoretical bases from which descriptions of school motivation have emerged, each one limited to one or two main explanatory constructs. Ball asked for a general theoretical integration of these separate descriptions of motivation, and cited a study by Chiu (1967) as a first move in that direction.

Chiu carried out a factor analytic study to investigate " the conceptual and empirical aspects of academic motivation ". Using samples of eleventh grade high-school pupils, Chiu administered to boys and girls separate inventories of some 250 items each. The items were drawn from a wide range of existing motivational and personality scales which operationalised most of the main motivational constructs described by Ball in relation to school motivation. The inventory, after item analysis, contained 16 motivational sub-scales which were factor analysed by the principal-axes method, followed by normalised varimax rotation to simple structure, A very similar structure was obtained for both boys and girls. It contained five motivational factors: positive orientation to school learning; need for social recognition; curiosity; motive to avoid failure; and reaction to expectation or conformity.

This study by Chiu represents the most thorough attempt to date to integrate the confusing welter of idiosyncratic constructs, but does not go as far as the ' general theoretical integration ' demanded by Ball. It does, however, provide strong evidence for the need to separate school motivation into at least five component parts.

An integrated approach to describing school motivation

In fact an integrated approach to school motivation had already been reported in considerable detail (Kozéki, 1975), but in Hungarian. Over a 10-year period, Kozéki has been attempting to identify a set of motivational dimensions which has a sound theoretical rationale, but which is also firmly rooted in the everyday experiences of pupils, teachers, and parents. To date over 1,000 interviews have been carried out and analysed to identify recurring descriptions of the sources of motivated behaviour. Rigorous qualitative analyses of these interview data established nine categories which spanned three separate motivational domains—affective, cognitive, and moral. Not only does this work establish an integrated framework to describe school motivation, but it also avoids the exclusive reliance on cognitive motives which was so marked in previous research in this area.

The nine motivational categories have been carefully justified in relation both to previous theories of motivation and to analyses of the extensive interview data (Kozéki, 1975). The underlying rationale of the scheme depends on the fact that children experience, in relation to learning, differing types of rewards and punishments. These depend essentially on relationships with others (the affective domain), developing competences in knowledge and skill (the cognitive domain), and the feelings of satisfaction derived from living up to the expectations of self and others (the moral domain). Within each of these domains Kozéki has been able to identify from the interview data three separable categories. As a description of these motivational categories has been published elsewhere in English (Kozéki, 1984), they are simply presented here in tabular form (Table 1).

TABLE 1

KOZÉKI'S NINE MOTIVATIONAL CATEGORIES

Label	Description of main sources of motivation
Affective domain	
Warmth	Encouragement and interest actively shown by parents
Identification	Feeling empathy with adults and wanting to please them
Sociability	Enjoying collaborative work and activities with peers
Cognitive Domain	
Independence	Satisfaction from working things out without help from others
Competence	Rewards from a recognition of developing knowledge and skills
Interest	Enjoyment derived from ideas
Moral Domain	
Trust	Satisfaction from doing things thoroughly and well
Compliance	Preferring the security of behaving according to defined rules or norms
Responsibility	Accepting the consequences of actions and monitoring own behaviour accordingly

Development of an inventory

It is, of course, difficult to present the results of interviews in a way which makes clear both the rigorous analytic procedures involved in establishing the main descriptive categories and the evidential justification of the choice of that particular set of categories. It was therefore decided to develop an inventory which would operationalise the nine motives as a series of sub-scales. Items were written to cover behaviour and attitudes which were considered to be expressions of those motives. Each item was carefully checked with pupils of varying ages to ensure that it was appropriately colloquial, and that the meaning given to it by the pupils coincided with the intended meaning.

A 100-item inventory was subsequently assembled from the item pool, and has been given to large samples of pupils, aged between 8 and 20, in Hungary, Czechoslovakia, Rumania, and East Germany. Factor analyses of the items (Kozéki, 1981) suggested that, within the school setting, the nine motives could be simplified to six main factors. In the affective domain (A), there were two factors. *Warmth* (A1) described the quality of relationships with both parents and teachers, while *sociability* (A2) in relation to school work was associated with both adults and peers. In the cognitive domain (B), independence rarely produced a separable factor, leaving just *competence* (B1) and *interest* (B2). In the moral domain (C), *compliance* (C2) remained a strong independent factor but *responsibility* (C1) and 'trust' merged.

Subsequently a 60-item inventory, based on the most effective items, was developed with six sub-scales representing the factors described above. This article reports the further development and use of this inventory in Britain and Hungary.

METHOD

Inventory of reactions to school and school work

The 60-item Hungarian inventory was translated into English and discussed with groups of pupils. Alterations were made to ensure that the vocabulary was appropriate. The English version was then translated back into Hungarian and again checked in discussions with pupils.

Instructions in the inventory asked pupils to give their reactions to

> " comments made by children at school. As the comments are opinions based on personal experience, there can be no right or wrong answers. We are interested in what *you* think."

The items were presented in a format which required a response on a 5-point scale ranging from " close to your own opinion " to " couldn't agree with that comment at all ".

The responses to the completed inventories were transferred to the computer using values ranging from 4-0 for each item, giving maximum sub-scale scores of 40 (10 items). The internal reliability of the sub-scales is reported later (Table 2). Evidence of validity has been published previously, in the form of substantial correlations (around 0·75) with teachers' ratings of the different forms of motivation (Kozéki, 1980), and is discussed subsequently in relation to the present study.

Sample

Four schools in Britain and three schools in Hungary were chosen to reflect the range of pupil intakes typical of the two countries. The British sample (N = 365; 180 boys and 185 girls) was drawn first from two single-sex independent schools with the usual preponderance of children of high socio-economic status. The third school was a comprehensive school in a country town with a full ability range of pupils, but with somewhat above average socio-economic status. The final British school was from a tower-block city area with a high proportion of pupils with below average levels of educational attainment.

The three Hungarian schools were all comprehensive schools in Budapest, but their intake areas meant that a majority of pupils came respectively from families in professional, mixed, and skilled and unskilled manual occupations (N = 800; 390 boys and 410 girls).

Analyses

Although the English version of the inventory had been carefully produced and checked for equivalence of meaning, there must be a remaining doubt about the equivalence of scores derived from inventories presented in different languages. In addition, the British sample had been chosen mainly for the purposes of checking the hypothesised motives built into the inventory. While the full range of ability and socio-economic status was covered in the sample, the extremes were over-represented. Statistics derived from this sample cannot be considered representative of British secondary schools as a whole, although the means of the different types of school should provide a reasonable indication of the range of differences to be expected. However, whilst the correlations will be higher, the pattern of inter-correlations, and the factor analyses, should be little affected by the heterogeneity of the sample.

The Hungarian sample is restricted through being limited to schools in the capital, but the schools represent a typical cross-section of schools in Budapest.

In spite of the remaining doubts about the equivalence of the schools and the acknowledged limitations of the samples, it was decided to carry out exploratory analyses of the means from subgroups of British and Hungarian pupils. Tests of statistical significance were not considered justifiable, except in relation to sex differences. Profiles of scores across the three types of school were compared, and pupils' responses to individual items were also examined.

The main analyses were designed to explore the intercorrelations between the subscales in the inventory, and to examine the factor-structure of the items in Britain and Hungary.

RESULTS AND DISCUSSION

Descriptions of the sub-groups

As previous studies had suggested that only six subscales could be fully justified within this particular inventory, the means and standard deviations in Table 2, and the profiles of schools in Figure 1, are based on that scoring scheme. The interpretations of differences in means were, however, informed by an examination of the means from the nine subscales.

The main conclusion to be drawn from this exploratory description of the six motivational dimensions is that the Hungarian means are consistently higher than the British means, and that the range of means between the different types of school is greater in Britain than in Hungary. This differing range is, however, dependent on the profiles of just two working-class schools and must thus be treated with caution.

Where sex differences are consistent between the three types of school, more general conclusions may be justified. The profiles show girls to have higher overall levels of school motivation than boys, with the largest and most consistent differences in both countries being in warmth, sociability, and compliance (t-tests gave $P<0{\cdot}05$ for British children and $P<0{\cdot}001$ for the Hungarians).

In the Hungarian sample competence scores were consistently higher than interest, while the reverse was true in the British sample, although the differences were small. Hungarian girls had very much higher scores in both responsibility and compliance than any other group. There was little apparent difference in mean responsibility scores between the samples of British and Hungarian boys, but the Hungarians showed higher levels of compliance.

Responses to individual items

Table 3 provides illustrations of the items used to define the nine hypothesised motives in the inventory, and also indicates the percentage agreement with these items of boys and girls in each country.

Although the caveat about the samples must be repeated, the responses to items about school and schoolwork are by no means as negative as recent studies on adolescents might have led us to anticipate (e.g., Gow and McPherson, 1980). In fact for 12-13-year-olds, attitudes to school do still seem to be favourable (Smith, 1980).

Table 3 shows a series of items in which pupils in the British sample are being more critical of teachers than those in Hungary (Items 2, 3, 4, 11, 12, and 18). The Hungarian children see collaborative work as more important than competition (Item 5), and see punishment in schools as fair (18). (There is no corporal punishment at all in Hungary.) In Britain, there was a higher recognition of the value of schoolwork (10), in spite of a more critical attitude to teachers. Other items show a remarkably similar pattern of responses between the two countries. These seem mainly to

TABLE 2

MEANS AND STANDARD DEVIATIONS OF SUBSCALE SCORES BY COUNTRY, SCHOOL, AND SEX—ALSO ESTIMATES OF SUBSCALE RELIABILITIES

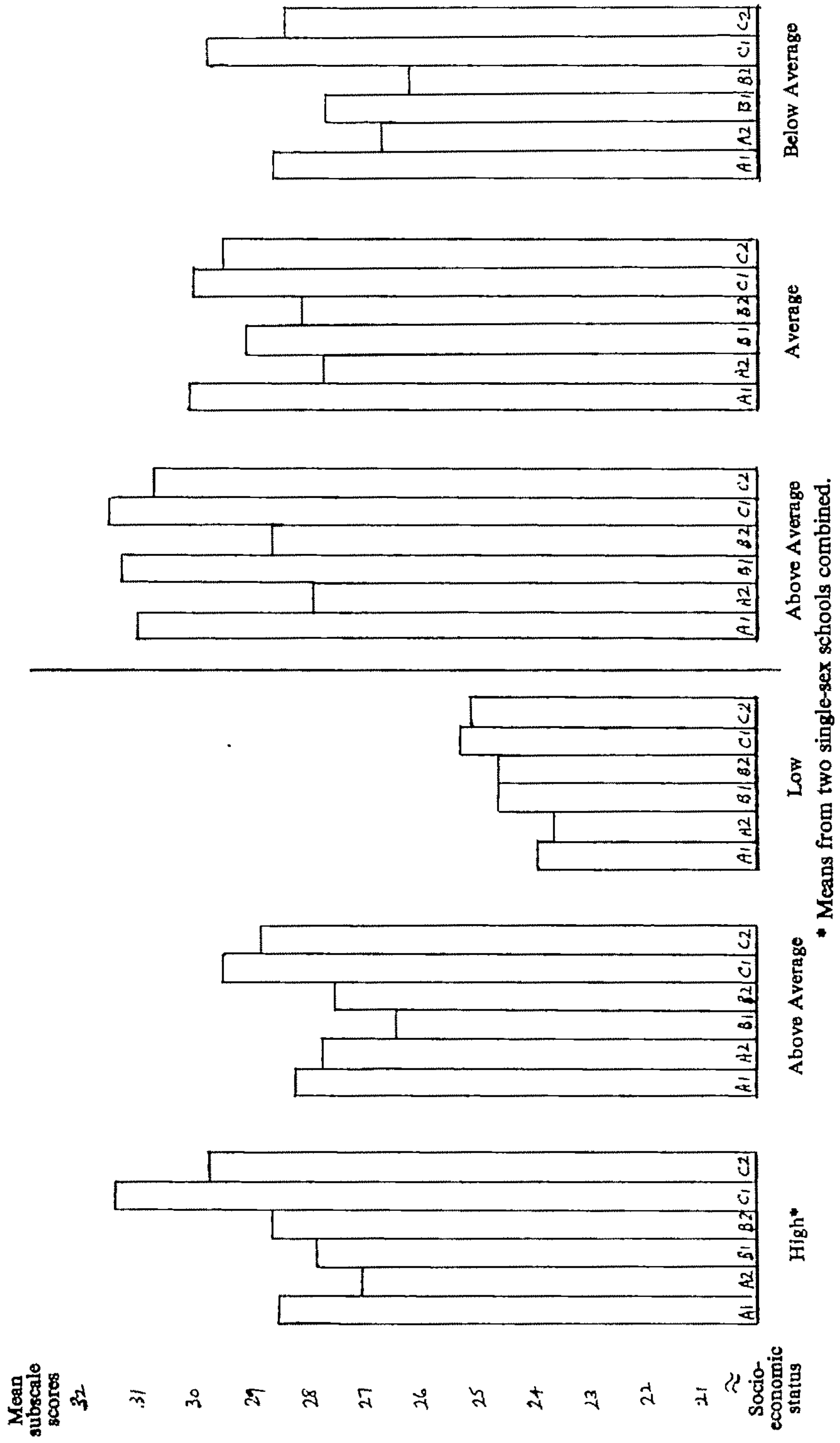
		Britain								Hungary								
Sub-scale		High SES B.	High SES G.	Above Av. B.	Above Av. G.	Low SES B.	Low SES G.	Total B.	Total G.	Above Av. B.	Above Av. G.	Average B.	Average G.	Below Av. B.	Below Av. G.	Total B.	Total G.	Estimate of Reliability*
(A1) Warmth	Mean	27·3	29·6	28·3	28·0	23·2	24·6	26·3	27·6	29·8	32·1	29·0	31·1	27·7	29·5	28·9	30·9	0·65 (B)
	SD	(5·4)	(4·2)	(5·7)	(5·4)	(5·6)	(5·4)	(5·9)	(5·4)	(4·9)	(4·4)	(5·6)	(4·8)	(4·7)	(4·7)	(5·2)	(4·8)	0 67 (H)
(A2) Sociability	Mean	26·0	27·9	27·5	27·9	22·3	24·9	25·3	26·9	26·2	29·5	27 0	28·4	26·6	26·8	26·7	28·2	0·64 (B)
	SD	(6·8)	(4·8)	(4·9)	(5·5)	(4·4)	(5·2)	(6·0)	(5·2)	(5·5)	(4·6)	(5·6)	(5·0)	(4·6)	(4·7)	(5·3)	(4·9)	0·68 (H)
(B1) Competence	Mean	27·8	27·8	26·7	26·1	24·0	25·2	26·4	26·6	30·4	32·1	28·2	29·9	27·4	28·0	28·7	29·9	0·55 (B)
	SD	(5·6)	(4·9)	(5·6)	(5·9)	(4·6)	(5·7)	(5·5)	(5·5)	(4·5)	(4·4)	(5·8)	(5·5)	(5·2)	(5·1)	(5·4)	(5·3)	0·67 (H)
(B2) Interest	Mean	28·3	28·8	27·8	27·1	24·1	25·1	26·9	27·2	27·5	29·7	27·5	28·6	25·8	26·6	27·0	28·2	0·51 (B)
	SD	(5·0)	(4·3)	(4·6)	(5·7)	(4·9)	(5·2)	(5·2)	(5·1)	(5·6)	(4·5)	(5·6)	(5·8)	(4·8)	(4·8)	(5·5)	(5·3)	0·67 (H)
(C1) Responsibility	Mean	31·4	31·4	29·9	29·0	24·5	26·0	28·9	29·1	30·8	32·1	28·7	31·3	29·1	30·4	29·5	31·2	0·72 (B)
	SD	(5·2)	(4·6)	(5·2)	(6·0)	(5·2)	(5·4)	(6·0)	(5·7)	(5·2)	(4·7)	(6·2)	(5·2)	(4·9)	(5·0)	(5·6)	(5·0)	0·74 (H)
(C2) Compliance	Mean	29·4	29·9	28·5	29·1	23·3	26·8	27·3	28·7	29·7	31·6	28·4	30·6	27·7	29·0	28·6	30·4	0·76 (B)
	SD	(6·4)	(5·2)	(6·5)	(6·9)	(6·1)	(7·0)	(6·9)	(6·3)	(5·5)	(4·7)	(6·2)	(5·1)	(4·8)	(5·1)	(5·7)	(5·1)	0·73 (H)
	N	76	85	50	39	54	61	180	185	124	127	166	147	100	136	390	410	365 (B) 800 (H)

Means and SDs to one place of decimals only. * Cronbach alpha.

FIGURE 1

PROFILES OF THE MEAN SUBSCALE SCORES OF SCHOOL MOTIVATION BY SCHOOL

Key: A1 = Warmth, A2 = Sociability, B1 = Competence, B2 = Interest, C1 = Responsibility, C2 = Compliance



* Means from two single-sex schools combined.

TABLE 3

PERCENTAGE AGREEMENT BY DIFFERENT SUB-GROUPS WITH SELECTED ITEMS

Motive	Items*		British			Hungarian	
		All	Boys	Girls	All	Boys	Girls
Warmth	1. My parents don't seem to be interested in whether I have done well in school (−)	9	10	8	10	10	9
	2. Nobody is really interested in trying to understand a child's feelings (−)	**30**	31	28	12	14	8
Identification	3. It would make me sad if I disappointed the teacher	34	36	32	**61**	55	**68**
	4. Most teachers never bother to explain well enough how to do things (−)	**37**	37	38	19	21	18
Sociability	5. Learning to work with other people is more important than trying to do better than them	67	60	**75**	**81**	75	**86**
	6. I am often in trouble in school (−)	18	**25**	10	12	**17**	8
Independence	7. Parents expect too much from children and give little help in return (−)	**19**	19	20	10	13	6
	8. I like to solve the problem myself rather than be told how to do it	59	57	62	57	56	57
Competence	9. I become so absorbed in some topics at school that I don't want to go on to other things	38	35	42	**49**	46	53
	10. School provides a great deal of knowledge about life	**79**	79	78	62	63	60
Interest	11. School is a boring place (−)	**33**	30	35	17	20	13
	12. Most school lessons are pretty dull and uninteresting (−)	**28**	27	**30**	18	**23**	14
Trust	13. When I don't do well at school I feel ashamed of myself and try to do bettter	**79**	70	**83**	68	64	**72**
	14. I would rather admit something I have done wrong than try to cover it up	67	62	**72**	69	63	**74**
Compliance	15. Pupils should support the aims of the school by obeying the rules	74	68	**81**	76	73	**78**
	16. Pupils should try to do what they are expected to do at school	74	70	**77**	75	72	**78**
Responsibility	17. I like to feel responsible for my actions whatever happens afterwards	58	**62**	53	**67**	60	**74**
	18. Punishment in schools is always unfair (−)	**31**	33	29	7	9	4
	(N)	(365)	(180)	(185)	(800)	(390)	(410)

Notes 1. * −ve sign indicates reverse scoring of the item.
2. Main types of differences indicated by bold type.

be items which are related more to psychological attributes than to educational or social milieux (8, 14, 15, 16).

Intercorrelations between the subscales

Table 4 presents the interrelationships between the six subscales separately for British and Hungarian samples. Among the six subscales, the highest intercorrelations are found between competence and compliance in both countries, and between

TABLE 4

Inter-Correlations with Subscales Separately by Country (N = 365, 800)

Subscale	(A1) Warmth		A2		B1		B2		C1	
	Brit.	Hung.	Brit.	Hung.	Brit.	Hung.	Brit.	Hung.	Brit.	Hung.
(A2) Sociability	69	63								
(B1) Competence	62	66	63	62						
(B2) Interest	58	58	55	62	52	71				
(C1) Responsibility	60	60	63	57	56	61	50	57		
(C2) Compliance	68	63	69	66	67	71	62	68	62	68

Decimal places omitted.

sociability (including identification with the teacher) and compliance, again in both countries. In Hungary there are higher correlations between interest and both competence and compliance, while in Britain higher correlations are found between warmth (in adult relationships) and both sociability and compliance.

The median values of correlations with each scale are within 0·01 between the two samples in four of the scales, with somewhat higher medians in the Hungarian sample for the two cognitive scales (competence and interest).

Overall the correlational matrices show strong similarities between the two samples, with correlations in the expected range of 0·5 to 0·75 implying the intended general factor of school motivation, but with sufficient differences between subscales to justify their retention as indicators of different aspects of school motivation. To find which aspects could be empirically justified, repeated factor analyses of the items were carried out.

Item analyses

Evidence for the existence of separate dimensions of school motivation was sought firstly by examining the correlations between items and adjusted scale totals, and subsequently by factor analyses.

In the Hungarian version the correspondence between items and hypothesised scales was remarkable with every one of the 54 items which related to the nine motives showing the highest correlation with the intended scale, with all but five having correlations above 0·5. In the English version the correspondence was less complete, but still substantial. There was, however, clear evidence of items where the translation, though linguistically accurate, had failed to coincide with the conceptual intention.

The interpretation of the factor analyses created more difficulty. As the items are all drawn from the same overall domain describing school motivation, and the reliability of individual items is inevitably low, tightly defined and clearly separable factors cannot be expected. It is also difficult to decide the best number of factors to extract. The SPSS program (Nie *et al.*, 1975) was used with alpha factor analysis (appropriate for scale construction) and oblique rotation (with delta set at zero). In a factor-analysis of inventory items the criterion based on eigen values above unity produces a large number of uninterpretable factors. And in a scree test from data of this type, the discontinuities rarely provide a clear indication of a single ‘best’ solution. It was therefore decided to combine the evidence from the scree test (which suggested between six and 10 factors) with the hypothesis-testing function of factor analysis (six or nine dimensions/categories could be predicted). Ten-, 8-, and 6-factor solutions were examined with the 8-factor solution producing the best balance between separation and interpretability of the groups of items defining the factors.

As all 60 items were drawn from a single measurement domain, factors precisely coinciding item-by-item with the subscales were not anticipated. It thus seemed crucial to examine separately the factor structure for boys, girls, and total sample in each country. Examination of the items in each factor with loadings above 0·4 produced a descriptive label, together with a note of items which described relationships with parents or teachers. Discussion of the factor descriptions from all six analyses led to agreement that there were essentially eight distinguishable factors which are shown in Table 5. The eight factors were, in descriptive terms at least, similar to the nine hypothesised motives, but in the various analyses they tended to merge to create between five and seven interpretable factors with also considerable variation in the specific defining items. Although the factors tended not to contain exactly the same set of items in each analysis, the meanings of the combination of items were identifiably similar across analyses—hence the possibility of restricting the number of descriptive labels to eight.

TABLE 5

SUMMARY OF THE FACTOR ANALYSES OF INVENTORY ITEMS

Hypothesised Motive	Scale	Description of Factor	Sub-Groups in which Factor Identified: British Total	British Boys	British Girls	Hungarian Total	Hungarian Boys	Hungarian Girls
Warmth	(A1)	Warm relationships with adults	*		*	*	*	*
Identification	(A2)	Good relationships with teachers and acceptance of school	*	*	*			
Sociability		Co-operative relationships with peers				**	*	*
Independence		Self-confidence and independence		*	**	*		
Competence	(B1)	Intention to work hard and seek knowledge	*	*	*		*	*
Interest	(B2)	Satisfaction with school and interest in school work	*			*	**	*
Trust		—						
Compliance	(C2)	Compliance with authority	*	*	*	*	**	*
Responsibility	(C1)	Acceptance of responsibility	*	*	*			
		(N)	(365)	(180)	(185)	(800)	(390)	(410)
Percentage of the variance explained by eight factors			43·1	46·9	45·4	41·3	43·1	43·7

Note: In some analyses a single factor split into two parts (**), otherwise the asterisk indicates the presence of that factor.

Combining the evidence from the item-scale correlations with the factor analysis, there appears to be reasonable justification for scoring the inventory on either six or nine subscales and for arguing that a similar factor structure of the inventory has been found in Hungary and Britain.

Although the hypothesised motives were separable, some of the overlaps between the factors were of particular interest. For example, although relationships between children and parents or teachers were described mainly by the first two factors, items relating to both of these dimensions also appeared in several other factors. The associations indicated in Table 6 were found in these analyses.

TABLE 6

ITEMS DESCRIBING PERCEPTIONS OF PARENTS AND TEACHERS ASSOCIATED WITH THE OTHER MAIN FACTORS

Motive	Pupil	Parents	Teachers
Independence	Self-confidence and independence	Pleased by good work, valuing opinions, and giving freedom	Providing help
Competence	Working hard and seeking knowledge	Perceived as interested and encouraging	Perceived as providing help
Interest	Satisfaction with school and interest in school work	Understanding, interested, and not over-demanding	Fair and providing help and good explanations
Moral Domain	Responsibility and compliance	Interested, trusting, supporting and encouraging	Fair and providing help

To illustrate these apparent links between attitudes to school and perceptions of treatment by parents and teachers, the items defining a factor of ' helplessness and lack of self-confidence in school ' (negative M5) among British boys are listed below in order of decreasing factor loadings (all above 0·30).

Items related to school and school work
I am often in trouble at school
When school work is hard I usually give up
I quickly lose interest if new topics are difficult
I am often told off at school
Much of what we learn at school is of no use to me
I always do my homework at the last minute

Items related to parents and the adult world
Nobody is really interested in trying to understand a child's feelings
Parents expect too much from children and give little help in return
One should not expect good ideas from pupils until they grow up
My parents don't seem to be interested in whether I have done well at school

Items related to teachers and school rules
Punishment in schools is always unfair
Most teachers never bother to explain well enough how to do things
It is usually the teacher's fault when you get into trouble at school
Pupils should be allowed to leave school at whatever age they want

In considering possible reasons for ' helplessness ' it must be recognised that other children in the same classes perceive the teachers as being fair and providing good explanations. There is presumably a substantial amount of transference from previous experiences involved in the attitudes to teachers, at least.

CONCLUSIONS

The major argument developed in this article is that there are several distinguishable categories within school motivation. The question to be answered is whether there is substantial evidence both for the psychological reality of several components and for the separability of as many as six or nine dimensions.

The high intercorrelations between the six inventory scales implies a strong single factor of school motivation. That conclusion has to be accepted, and indeed the total score on the inventory could be used as a good indicator of a pupil's overall school motivation. But the item-scale correlations also provide a clear justification, at least in the Hungarian version, for retaining the six hypothesised scales. In addition the emergence of between five and seven interpretable factors in each of six factor analyses, and the identifiable correspondence between the factor structures from two such markedly different educational and social systems, provides good grounds for accepting at least internal validity for the six dimensions.

The argument for the psychological reality, or external validity, of the various dimensions draws again on several threads. There is evidence external to the inventory itself, but related to it, which was described earlier. The extensive interviews and the theoretical rationale which led to the description of nine motives, together with evidence from correlations with teachers' ratings, underpin the evidence derived from the item analyses. The factor analyses represent, at the least, an independent categorisation of the motive structure which produces a broadly similar pattern.

Details of the interview and other analyses are available only in Hungarian (Kozéki, 1975, 1980), and therefore have to be accepted on trust here. But the evidence for this form of external validity is none the less real and substantial. Other evidence of external validity can be found in the similarities between the descriptions of the factors reported here and those described by Chiu in the most extensive of the previous analyses of motivational structure. The factor descriptions are set out below (Table 7) in a way which is intended to indicate the correspondences, although

TABLE 7

Descriptive Comparison of Two Sets of Factor Analyses of School Motivation

Chiu	Kozéki and Entwistle
Affective	
	Warm relationships with adults
Motive to avoid failure	
	Good relationships with teachers and acceptance of school
Need for social recognition	
	Co-operative relationships with peers
Cognitive	
	Self-confidence and independence
Positive orientation towards school learning	
	Intention to work hard and seek knowledge
Curiosity	
	Satisfaction with school and interest in school work
Moral	
	Acceptance of responsibility
Conformity	
	Compliance with authority

of course a further empirical study would be required to confirm these indicative connections.

There is thus a growing body of evidence available to justify the separation of school motivation into five or more dimensions, and some agreement as to how those dimensions are to be described. It is suggested that the particular formulation used in this study has two particular strengths. It is based on a theoretical analysis of the sources of satisfaction which are utilised by different children, and it emphasises equally the affective/social, cognitive, and moral domains. The explanatory power of the concept of school motivation is considerably strengthened by such a separation into its component parts. Moreover, at a time when educationalists are increasingly criticising the over-emphasis in schools of cognitive objectives, the availability of an inventory covering a wider spectrum of pupils' behaviours and attitudes is particularly valuable.

Although this study has concentrated on establishing the factor structure of the inventory, it has pointed up two important additional uses for it. First the use of school (or class) profiles (as in Figure 1) might prove to be an interesting way of describing certain aspects of school ethos (or class ' personality '). Then the suggestive links between perceived parental treatment and pupils' attitudes towards school and schoolwork would be well worth exploring further. Although these links can only be considered tentative on the basis of the analyses presented here, there already exists substantial evidence based on interviews and questionnaires in Hungary (Kozéki, 1976) that the satisfaction pupils derive from school, and their attitudes to teachers, are firmly rooted in their relationships with their parents. This would seem to be an important, and timely, area for further research.

ACKNOWLEDGMENTS.—We are grateful to Alastair Pollitt for carrying out the statistical analysis, and to Willie Bremner, Eric Millington, and Ian Selmes for collecting the British data.

REFERENCES

ATKINSON, J. W., and FEATHER, N. T. (1966). *A Theory of Achievement Motivation.* New York: Wiley.

BALL, S. (1977). *Motivation in Education.* New York: Academic Press.

BIRNEY, R. C., BURDICK, H., and TEEVAN, R. C. (1969). *Fear of Failure.* New York: Van Nostrand.

CATTELL, R. B. (1957). *Personality and Motivation Structure and Measurement.* New York: World Books.

CATTELL, R. B., and CHILD, D. (1975). *Motivation and Dynamic Structure.* New York: Wiley.

CHIU, L. H. (1967). A factorial study of academic motivation. Unpublished doctoral dissertation, Teachers College, Columbia University.

ENTWISTLE, N. J. (1968). Academic motivation and school attainment. *Br. J. educ. Psychol.*, 38, 181-188.

ENTWISTLE, N. J., and RAMSDEN, P. (1983). *Understanding Student Learning.* London: Croom Helm.

ENTWISTLE, N. J., THOMPSON, J. B., and WILSON, J. D. (1974). Motivation and study habits. *Higher Educ.*, 3, 379-396.

ENTWISTLE, N. J., and WILSON, J. D. (1977). *Degrees of Excellence: the Academic Achievement Game.* London: Hodder and Stoughton.

FINGER, J. A., and SCHLESSER, G. E. (1965). Non-intellective predictors of academic success in school and college. *School Rev.*, 73, 14-29.

GOW, L., and MCPHERSON, A. (1980). *Tell Them From Me.* Aberdeen: Aberdeen University Press.

JONES, C. L., MACKINTOSH, H., and MCPHERSON, A. F. (1973). Questions of uncertainty: non-cognitive predictors of achievement. In PAGE, C. F., and GIBSON, J. (Eds.), *Motivation: Non-Cognitive Aspects of Student Performance.* London: SRHE.

KOZÉKI, B. (1975). *Motivalas es Motivacio* (Motivating and Motivation). Budapest: Tankonyukiado.

KOZÉKI, B. (1976). Az iskolai tanulas motivaciojanak vizsgalata (Investigation of motivation to study). *Magyar Pszichologiai Szemle*, 33, 335-346.

KOZÉKI, B. (1980). *A Motivalas es Motivacio. Osszefuggeseinek Pedagogiai Pszichologiai vizsgalata* (Pedagogical Psychological Examination of Relationship between Motivating and Motivation). Budapest: Akademiai Kiado.
KOZÉKI, B. (1981). Aziskolai teljesitmeny dinamikus faktorai (Dynamic factors determining academic achievement). *Magyar Pedagogia*, 25, 358-367.
KOZÉKI, B. (1984). Motives and motivational styles in education. In ENTWISTLE, N. J. (Ed.), *New Directions in Educational Psychology*. Brighton: Falmer Press (in press).
KOZÉKI, B., and ENTWISTLE, N. J. (1983). Describing and utilising motivational styles in education. *Br. J. educ. Stud.*, XXXI, 184-197.
MCCLELLAND, D, C., ATKINSON, J. W., CLARK, R. A., and LOWELL, E. L. (1953). *The Achievement Motive*. New York: Appleton-Century-Crofts.
NIE, N. H. *et al.* (1975). *Statistical Package for the Social Sciences.* New York: McGraw Hill.
SMITH, D. J. (1980). Opinions of school, academic motivation and school adjustment in the first year of secondary education: a pilot study in West Yorkshire. *Educ. Stud.*, 7, 177-183.

(*Manuscript received 22nd February*, 1984)

Br. J. educ. Psychol., 54, 320-330, 1984

THE BEHAVIOUR OF PERSISTENT SCHOOL ABSENTEES

By K. REID

(*West Glamorgan Institute of Higher Education, Swansea*)

Summary. The Rutter (1967) Children's Behaviour Questionnaire (Scale B) was given to 77 persistent absentees and their two control groups in the third, fourth and fifth years at a large inner-city comprehensive school in a deprived area in South Wales over a three-week period. The results showed that the persistent absentees depicted significantly worse behavioural traits than the pupils in the two control groups and exuded higher levels of neurotic and anti-social conduct. Some interesting group and sex differences were found within these data. The implications of these results are discussed in the light of recent findings.

INTRODUCTION

Until the late 1970s there was an imbalance in the amount of research undertaken into school absenteeism. Even now, much more is known about some of the home background (Tyerman, 1968), social (May, 1975) and psychological (Reid, 1982a) than institutional (Reynolds *et al.*, 1976; Rutter *et al.*, 1979) processes involved in this phenomenon (Reid and Kendall, 1982).

The family and home circumstances of absentees and truants are frequently characterised by conditions of multiple deprivation (Carroll, 1977) and low social class (Davie *et al.*, 1972; Fogelman and Richardson, 1974). The adverse home and social conditions of absentees and truants include large families and overcrowding (Brooks *et al.*, 1962; May, 1975; Tibbenham, 1977), inconsistent home discipline and father absence (Hersov, 1960), marital stress (Trigg, 1973), free school meals (Galloway, 1976), one parent families (Elliott, 1975; Ferri, 1976), death of a parent (Watkin, 1975), and unemployment and illegitimacy (May, 1975).

Recent research has suggested that the home and social backgrounds of absentees and truants need to be investigated in conjunction with their psychological and institutional aspects. For example, reasons for absenteeism given by pupils indicate the importance of interaction between family and social factors with school experiences (Buist, 1980; ILEA, 1981). It is becoming increasingly apparent that persistent absentees tend to blame their schools rather than themselves or their home backgrounds for their non-attendance (Reid, 1983a). Indeed, the number of reasons given by persistent absentees to justify their behaviour differs and increases between the initial and chronic stage (Reid, 1983b).

The question of disruptive behaviour in school being related to persistent school absenteeism is not new. During the last 15 years there have been a number of reports on disruptive behaviour which have been linked to absenteeism (NAT, 1975; UCAC, 1975; EIS, 1975; Pack Report, 1977; HMI, 1978; NAS/UWT, 1981; Evans, 1981). The HMI Report, for example, surveyed 18 schools and noted the general incidence of behavioural problems and absenteeism. It went on to examine ways in which schools are dealing with these issues through their pastoral care provision, inter-agency liaison, school and legal sanctions, and off- and on-site units.

Similarly, there has been a great deal of research into disruptive behaviour in schools (Caspari, 1976; Jones-Davies and Cave, 1976; Laslett, 1977; Lawrence *et al.*, 1977; Bird *et al.*, 1980; Tattum, 1982) and its relationship to absenteeism/truancy. Illustratively, Bird *et al.* (1980) looked at six schools in two outer London boroughs and investigated the link between pupil motivation and disaffected behaviour such as truancy and disruption. They used material obtained from case studies to analyse the pupils' experiences at home, in the community and in employment. Like Galloway *et al.* (1982), they found that different models of pastoral care were adopted by the schools and differential use made of alternative education units.

The relationship between disruptive school behaviour and problems like school absenteeism led Davie (1980) to obtain a sizeable grant from the Welsh Office in order to organise a unique course and research project in South Wales (where there is a higher and disproportionate amount of absenteeism: Davie *et al.*, 1972; Carroll, 1977; Reynolds, 1982) for senior and middle management staff from many secondary schools. The concept behind this course is to enable teachers to become better informed and equipped to cope with the problems of disruptive behaviour and absenteeism in their schools by studying for an initial/higher degree. After qualifying, it is envisaged that these teachers will lead the ' fight ' against disruptive conduct and absenteeism in their institutions.

A few accounts exist of the way in which ordinary secondary schools have tried to overcome their own absenteeism problem (Jones, 1980; Reid, 1982b). Much more, however, has been written about how units in ordinary and ' special ' schools cope with their chronic truancy cases (Grunsell, 1978). Overcoming behavioural problems related to absenteeism in ordinary schools is extremely hard work. In special units and schools some limited success has been reported. This form of ' treatment ' rarely leads to a truant returning to his original school and being successfully reintegrated. Consequently, not all teachers are in favour of these schemes. Some teachers argue that specialist units ' normalise ' deviant behaviour and decrease the commitment of form and subject teachers to tackle the problems for themselves.

Finally, it should be noted that the title uses the term ' persistent school absenteeism ' rather than truancy. This choice of words is used because it encompasses the varieties of absences from school and does not carry with it the connotations commonly associated with other labels such as truancy (Carroll, 1977).

METHOD

The samples

Three samples were chosen in order to investigate the relationship between persistent school absenteeism and the pupil's behaviour. The Persistent School Absentees were selected by senior staff in the schools and the education welfare officers and were known to have missed school for a minimum of 65 per cent or more of school time in the year preceding the study. The vast majority, however, had been categorised as being Persistent Absentees for about two or three years with some 25 per cent of them having had serious attendance problems in their primary schools.

The Control Group 1 pupils came from the same forms as the Persistent Absentees and were matched for age and sex. The Control Group 2 pupils originated in the highest forms in the school and like the Control Group 1 pupils, were matched by age and sex to the nearest month with the Absentees (cf. Stott, 1966). All the pupils in the two control groups were consistently good attenders, usually recording on average 100 per cent attendance or close to it during each school term. All the pupils in the three groups came from the third, fourth, and fifth years in the school in order to maintain a manageable sample given the nature of the problem. The three groups were chosen in order to maximise the possible differences. As absentees and good attenders select themselves, a randomised design was not possible.

The research was sub-divided and undertaken in stages. In stage one, the Persistent Absentees and good attenders were selected and matched. The second part of the enquiry investigated and compared the social and educational backgrounds of the pupils (Reid, 1981, 1982a) using a social anthropological approach (Stubbs and Delamont, 1976). In the third stage, the Rutter (1967) Children's Behaviour Questionnaire was used.

Measures used in the study

Rutter devised the Children's Behaviour Questionnaire for analysing pupils' behaviour. For the purposes of this study, Scale B was used. This was completed by the form teachers of each of the 77 pupils in the three groups under instructions given by the researcher over a three-week period towards the end of the investigation. The instrument was not only used to measure the pupils' behaviour but also to ascertain their scores for neuroticism and anti-social behaviour.

While the questionnaire was primarily designed for use with 7- to 13-year-old children, it has been used successfully with older pupils, notably on samples of maladjusted subjects. It should be noted, however, that Rutter himself outlined several important limitations on the use of his scale. These can be found by referring to the instructions. Rutter has also reported test-retest reliability coefficients of 0·89 (interval two months) and 0·85 (interval 13 weeks), and an inter-rater reliability coefficient of 0·72.

Three null hypotheses were formulated prior to the use of the scale: that no significant differences would be found between the three samples on main scale, neuroticism, and anti-social behaviour.

Analysis of the data

In order to interpret the data, analysis of variance and t-tests were used to ascertain whether the reported behaviour of the Persistent Absentees was worse than for the two control groups. Two-way analysis of variance scores by group and sex were also used when investigating the intra- and inter-group differences. The data were analysed mechanically, using SPSS version 8 (Nie *et al.*, 1975). Three levels of probability were considered significant as shown in the tables. A pilot study was organised at a nearby comprehensive school before the main work was undertaken.

Constraints

The main constraints lay in the fact that some of the form teachers knew their Persistent Absentees less well than the attenders. Fortunately, due to the organisation of the school, this proved to be less of a deterrent than might otherwise have been the case.

The corollary of this argument also applied. Some of the form teachers had a greater knowledge and were more aware of the individual circumstances of the Persistent Absentees as these deviant pupils took up far more of their time in some ways than the regular attenders. Support for this latter argument was found on the completed questionnaires, where form teachers recorded a much higher ratio of additional comments in the space provided on the Absentees than the controls. The co-operation of the school staff in the enterprise was such that a 100 per cent return of the questionnaires was achieved.

RESULTS

Tables 1, 3, and 5 show the distribution of the overall results for the Persistent Absentees and the controls on the Rutter Main Scale and its two subscales respectively. Tables 2, 4, and 6 compare the results between the groups on the three measures.

Table 1 reveals that the mean for the Persistent Absentees was three and four times greater than those for Control Groups 1 and 2 respectively. Not surprisingly, two variables—" truants from school " and " tends to be absent from school for trivial reasons "—depicted the greatest differences between the Persistent Absentees and the two control groups.

TABLE 1

Distribution of the Overall Results for the Absentees and Attenders on the Main Scale

Group	N	M	SD	F Ratio	P
Absentees	77	7·26	4·48	53·54	0·001
Control Group 1	77	2·10	3·21		
Control Group 2	77	1·74	3·27		

When these data were re-analysed and the two criterion variables above were excluded, the differences between the three groups were much reduced. The differences between the academic controls and the Absentees, however, remained greater than those for the Absentees and their matching controls from the same forms.

Further large differences (nine or more points between the groups) were obtained on seven items. These were: " not much liked by other children ", " tends to do things on his own—rather solitary ", " often appears miserable, unhappy, tearful or distressed ", " has poor concentration or short attention span ", " often tells lies ", " is often disobedient ", and " has stolen things on one or more occasions ".

TABLE 2

A Comparison of the Results Among the Three Groups on the Main Scale

Group	N	M	SD	Group	N	M	SD	t	P
Absentees	77	7·26	4·48	Control Group 1	77	2·10	3·21	8·2	0·001
Absentees	77	7·26	4·48	Control Group 2	77	1·74	3·27	8·7	0·001
Control Group 1	77	2·10	3·21	Control Group 2	77	1·74	3·27	0·7	

Table 2 shows that there were highly significant differences between the scores for the Persistent Absentees and the two control groups on the Rutter Main Scale. There were no statistically significant differences between the two control groups.

TABLE 3

Distribution of the Overall Results for the Three Groups on the Subscale for Neuroticism

Group	N	M	SD	F Ratio	P
Absentees	77	0·77	1·28	5·38	0·01
Control Group 1	77	0·27	0·72		
Control Group 2	77	0·32	1·01		

Table 3 shows that there were large and statistically significant differences among the overall results for the three groups. Interestingly, the mean for the Control Group 2 pupils was slightly higher than for those in Control Group 1. Presumably this is partially explained by the composition of the Control Group 2 sample as academic pupils are known to suffer from greater levels of anxiety than the ' normal ' population. Nevertheless, too much should not be made of this point as the mean for all the three groups was very low.

TABLE 4

A COMPARISON AMONG THE THREE GROUPS ON THE SUBSCALE FOR NEUROTICISM

Group	N	M	SD	Group	N	M	SD	t	P
Absentees	77	0·77	1·28	Control Group 1	77	0·27	0·72	2·9	0·01
Absentees	77	0·77	1·28	Control Group 2	77	0·32	1·01	2·4	0·05
Control Group 1	77	0·27	0·72	Control Group 2	77	0·32	1·01	0·4	

Table 4 shows that there were statistically significant differences between the scores for the Absentees and each of the two control groups on the subscale for neuroticism. There were no major differences between the two control groups.

Tables 3 and 4 reveal that the vast majority of the pupils in all three groups, including 61 per cent of the Absentees, showed no signs of any neurotic behaviour.

TABLE 5

DISTRIBUTION OF THE OVERALL RESULTS BETWEEN THE THREE GROUPS ON THE SUBSCALE FOR ANTI-SOCIAL BEHAVIOUR

Group	N	M	SD	F Ratio	P
Absentees	77	1·08	1·94	6·44	0·01
Control Group 1	77	0·38	1·06		
Control Group 2	77	0·34	1·15		

Table 5 shows that there were highly significant differences among the three groups on the subscale for anti-social behaviour. Once again, the means for the two control groups were very close.

TABLE 6

A COMPARISON OF THE RESULTS AMONG THE THREE GROUPS ON THE SUBSCALE FOR ANTI-SOCIAL BEHAVIOUR

Group	N	M	SD	Group	N	M	SD	t	P
Absentees	77	1·08	1·94	Control Group 1	77	0·38	1·06	2·8	0·01
Absentees	77	1·08	1·94	Control Group 2	77	0·34	1·15	2·9	0·01
Control Group 1	77	0·38	1·06	Control Group 2	77	0·34	1·15	0·2	

Table 6 reveals that there were statistically significant differences between the Persistent Absentees and each of the two control groups on the subscale for anti-social behaviour. As in Tables 3 and 4, the vast majority of the pupils in all three groups, including 65 per cent of the Absentees, showed no signs of any anti-social conduct in the estimation of their form teachers.

TABLE 7

A Comparison Between the Intra- and Inter-Group Differences on the Three Measures Using Two-Way Analysis of Variance

		MAIN EFFECTS							
		INTRA (SEX)				INTER (GROUP)			
Scale	Groups	df	M.Sq.	F	P	df	M.Sq.	F	P
Rutter	All	1	100·57	7·53	0·01	2	733·81	54·92	0·001
Rutter	Abs & CG1	1	112·03	7·65	0·01	1	1,023·44	69·86	0·001
Rutter	Abs & CG2	1	53·69	3·54	—	1	1,172·89	77·42	0·001
Rutter	CG1 & CG2	1	44·26	4·30	0·05	1	5·09	0·495	—
Neuroticism	All	1	0·07	0·07	—	2	5·62	5·55	0·01
Neuroticism	Abs & CG1	1	0·08	0 07	—	1	8·89	8·16	0·01
Neuroticism	Abs & CG2	1	0·009	0·007	—	1	7·96	6·34	0·05
Neuroticism	CG1 & CG2	1	0·08	0·112	—	1	0·03	0·04	—
Anti-Social	All	1	11·30	6·73	0·01	2	17·73	10·56	0·001
Anti-Social	Abs & CG1	1	10·61	4·91	0·05	1	26·59	12·31	0·001
Anti-Social	Abs & CG2	1	10·61	4·88	0·05	1	26·60	12·24	0·001
Anti-Social	CG1 & CG2	1	2·96	4·21	0·05	1	0·00	0·00	—

TABLE 8

A Comparison Between Male and Female Pupils by Group on the Rutter Main Scale and its Two Subscales

Scale	Group	Sex	N	M	Sex	N	M	df	t	P
Rutter	Absentees	M	44	8·04	F	33	6·21	75	1·81	0·05
Rutter	Control Group 1	M	44	2·79	F	33	6·21	75	1·81	0 05
Rutter	Control Group 2	M	44	1·99	F	33	1·42	75	0·74	0·05
Neuroticism	Absentees	M	44	0·77	F	33	0·76	75	0·05	—
Neuroticism	Control Group 1	M	44	0·29	F	33	0·20	75	0·32	—
Neuroticism	Control Group 2	M	44	0·32	F	33	0·33	75	0·06	—
Anti-Social	Absentees	M	44	1·41	F	33	0·64	75	1·80	0·05
Anti-Social	Control Group 1	M	44	0·55	F	33	0·15	75	1·70	0·05
Anti-Social	Control Group 2	M	44	0·43	F	33	0·21	75	0·84	0·05

TABLE 9

A Comparison of the Inter-Group Differences on the Three Measures

Scale	Sex	Group	N	M	Group	N	M	df	t	P
Rutter	M	Absentees	44	8·05	Control Group 1	44	2·73	86	6·01	0·001
Rutter	M	Absentees	44	8·05	Control Group 2	44	1·99	86	7·74	0·001
Rutter	F	Absentees	33	6·21	Control Group 1	33	1·18	64	5·95	0·001
Rutter	F	Absentees	33	6·21	Control Group 2	33	1·42	64	4·65	0·001
Neuroticism	M	Absentees	44	0·77	Control Group 1	44	0·29	86	2·18	0·05
Neuroticism	M	Absentees	44	0·77	Control Group 2	44	0·32	86	2·11	0·05
Neuroticism	F	Absentees	33	0·76	Control Group 1	33	0·24	64	1·97	0·05
Neuroticism	F	Absentees	33	0·76	Control Group 2	33	0·33	64	1·29	—
Anti-Social	M	Absentees	44	1·41	Control Group 1	44	0·54	86	2·25	0·05
Anti-Social	M	Absentees	44	1·41	Control Group 2	44	0·43	86	2·58	0·01
Anti-Social	F	Absentees	33	0·63	Control Group 1	33	0·15	64	1·78	0·05
Anti-Social	F	Absentees	33	0·63	Control Group 2	33	0·21	64	1·36	—

DISCUSSION

It is important to discuss the implications of the findings further both for their own sake and in the light of previous research. A closer inspection of the results revealed that the Absentees had some behavioural problems which were not wholly confined to their erratic attendance at school. It may be, therefore, that further research is needed into the personalities of school absentees as some of the conduct disorders (" is often disobedient ", " often appears miserable, unhappy, tearful or distressed ") are not explainable purely in terms of the non-attendance. One of the constraints on work of this kind, however, is that existing personality inventories and tests may not be sufficiently developed and sophisticated for the task (Billington, 1979).

These contentions are supported by the facts. For example, while there was some minority support for Tyerman's (1968) assertion that truants (note: not absentees) are frequently " lonely " and " unhappy " people with few friends, over three-fifths of the Absentees in this study did not depict any neurotic symptoms in the estimations of their form teachers. Furthermore, the evidence shows that even amongst hardened school absentees, anti-social conduct is a minority tendency. In this enquiry, disruptive behaviour was attributable to 35 per cent of the Absentees (Reid, 1983c). Interestingly, the findings reported elsewhere (Reid, 1982c) that the Control Group 1 pupils claimed to have more friends in school and in their own forms than the Persistent Absentees and Control Group 2 pupils suggests that some less able and potential disruptive lower ' stream ' students attend school for compensatory social rather than educational reasons. If so, this suggests that the temperamental and personality characteristics of pupils like truants and absentees are even more important than some people have hitherto realised. If not, why should so many pupils from low social class and deprived home backgrounds attend school when others do not? This could also partially explain why male and female absentees appear to have so much in common as it may be their personality types rather than sex *per se* which leads them to withdraw from school.

Tables 7, 8, and 9 present the intra- and inter-group differences on the three measures and make this latter contention clearer. Throughout most of these data, male pupils tended to score higher than their female counterparts. Illustratively, the overall means for the male pupils in all three groups were higher than those for the females on both the Rutter Scale and its anti-social subscale. By contrast, the means for the two sexes on the neuroticism subscale were almost identical, although interestingly, the females in Control Group 2 scored higher than the males on this measure. These intra-group differences reached significant proportions at the 5 per cent level on six of the nine calculations (see Table 8).

Likewise, Table 9 shows that there were statistically significant differences between the male Absentees and female good attenders on 10 of the 12 measures. Thus, there were large and significant differences between the male Absentees and male good attenders in both control groups on each of the three scales. There were also significant differences between the female Absentees and female Control Group 1 pupils on every measure. There were, however, statistically significant differences between the female Absentees and female Control Group 2 good attenders only on the Rutter Main Scale. Presumably this latter finding indicated that general behaviour rather than neuroticism or anti-social conduct was the main difference between the Absentees and their academic controls.

Hence, Table 8 shows that the means for the male and female Absentees were approximately three times as great as the good attenders in both control groups on the main scale and on most of the comparisons of the two subscales. This is an important result as it confirms that the behaviour of both male and female Absentees has more in common with one another than with good attenders.

Nevertheless, another fact should not be overlooked. The male Absentees tended to score more frequently on those items which related to either maladjusted or delinquent conduct than the girls. Once again, these findings accord well with previous studies (Tennent, 1971; West, 1982). Hence, although the total scores for the male and female Absentees were similar, there were considerable item-differences between them. This is only to be expected in view of well-known sex differences between adolescent male and female pupils aged 13 to 15-plus.

What the evidence seems to suggest is that some pupils are better able to cope with all forms of negative reinforcement—social, psychological, and institutional—than their peers, irrespective of whether they are male or female. If so, the importance of the pupils' self-concepts are again apparent. Furthermore, the precise correlation between the low self-concepts of Persistent Absentees (Reid, 1982a) and their behaviour in school is another area which requires further research.

Recent evidence has shown the substantial differences between schools in their attendance rates even in comparatively homogenous areas (Reynolds *et al.*, 1976; Rutter *et al.*, 1979). Although there have been too few action-based studies into the school processes involved in absenteeism, there is now some evidence which shows that the curriculum, school rules and teachers and their teaching styles are three important variables in the search (Reid, 1983a). Moreover, evidence obtained from the use of a repertory grid technique (Kelly, 1955) shows that Persistent Absentees tend to hold their parents in higher esteem than their teachers irrespective of their personal backgrounds (Reid, 1983d). Consequently, much more research is needed into the institutional factors and processes which are related to absenteeism and truancy.

Some work in this latter area has already been undertaken. For example, Galloway (1976) analysed data obtained from 30 comprehensive schools in Sheffield on children missing 50 per cent of possible attendances in an autumn term. He found that there was no association between size of school and persistent absenteeism. There was, however, an association between socio-economic hardship and absenteeism in the school catchment area. None of these factors was associated with suspension rates, but the former status of the comprehensive as a selective school was related to high rates of suspension.

The work of Galloway *et al.* (1978, 1982) in their Sheffield School and Home Project is important to our understanding of disruptive behaviour and absenteeism. The authors describe the provision of special groups within schools for disruptive pupils, their administration and curriculum, and teacher and pupil perceptions of their purpose and effectiveness. They conclude that, although all schools contain potentially disruptive pupils, some are clearly more able and willing to contain them than others. The researchers hint that certain schools only contain their difficult pupils at the expense of the conforming majority. Thus, exclusion and suspension rates vary enormously. In many cases, these depend upon individual teacher responses and which school the pupils happen to be attending. As things now stand, there are clear differences between schools in the way in which they operate their suspension and exclusion policies.

Likewise, Galloway *et al.* found that the instigation of legal sanctions to non-attenders and their parents lacks uniformity. It may well be that different policies and practices exist between local authorities when implementing their own suspension, exclusion or punishment procedures for disruptive conduct, but so far there is little evidence on this issue. Similarly, very little information is available on the stress imposed on teachers who come into frequent contact with disruptive pupils (Dunham, 1981). Galloway *et al.* discuss the reasons which contribute to teacher stress and the ways in which this is affected by differing models of discipline and pastoral care in schools. They recommend tackling problem behaviour through a schools policy, organisation and ethos by having clearly defined goals.

A fascinating insight into the analysis of the rules which operate within classrooms and the ways in which routine deviance is expressed by pupils and imputed by teachers is provided in the account of Hargreaves *et al.* (1975). This Manchester-based study describes the method of social typing used by teachers to define certain pupils as potential troublemakers and the strategies adopted by them to maintain control in the classroom. Their account highlights the importance of teacher personality and ability in reacting to potentially difficult situations in the classroom. The fact that it is very hard to ' teach ' classroom control on initial and in-service courses only reinforces the need for headteachers and middle management staff to keep a constant watch on potentially difficult situations in their schools.

The large-scale DES project into the Structure and Process of Initial Teacher Education in universities in England and Wales [SPITE] (Reid and Patrick, 1980; Patrick *et al.*, 1983) found that a majority of a cohort of post-graduate student teachers were inherently concerned about their ability to maintain order in the classroom at the start of their courses. The students felt that this was the one issue which could make or break them during their training. In the event, 49·2 per cent of the students admitted having some problems on teaching practice in being able to control difficult classes and 5·6 per cent admitted to major difficulties. Given the lack of help currently available for probationary (DES, 1982) and qualified teachers in schools, it is not clear how staff acquire their ability to cope in classrooms, especially with potentially disruptive situations. Evidence obtained from SPITE showed that student-teacher problems in maintaining control on teaching practice resulted from: " not being sufficiently firm at the start of teaching practice " (50·4 per cent), " having different standards of discipline from those of the regular teachers " (39·8 per cent), " attempting to be too friendly with the pupils " (38·9 per cent), " taking over classes from weak teachers " (29·6 per cent), " teaching topics/subjects which you were not sufficiently familiar" (28·9 per cent), " being inadequately prepared for the problems of discipline " (28·7 per cent), and " being unable to spot potentially difficult situations " (23·9 per cent).

The researchers also found clear sex, subject, regional, and institutional differences within these data. For example, female post-graduate student teachers reported having fewer disciplinary problems on teaching practice than males; science students (especially maths and physics students) stated that they had more difficulties than art, modern language, social science, primary and middle and physical education trainees.

Reid (1983c) has presented some case data and described some of the statements made by a few of the disruptive absentees in his sample. These exemplify the type of outbursts which some absentees claim can and do take place in schools. These data also show the various changes which occur in the estimations of teachers about the conduct of certain absentees over a period of time, based on information obtained from school reports and other documents.

Finally, some caution is necessary before the findings presented in this paper are generalised. It needs to be remembered that the form teachers' estimations of the absentees were influenced to some extent by their own experiences. When teachers consider a pupil to be deviant, they tend to attribute him with a host of related quirks (Hargreaves *et al.*, 1975).

CONCLUSIONS

The evidence presented in this paper suggests that there are fundamental differences between the behaviour of the Persistent Absentees and good attenders which are not solely confined to their attendance at school. In some ways, the behaviour of both male and female absentees has more in common with one another than with good attenders, irrespective of the latters' ability or social origins.

The findings reinforce the need for further research to be undertaken which investigates the interaction between the curriculum, teachers and their teaching styles, and school absenteeism. It is important that the conduct of school absentees is not seen as a subject in isolation. As absenteeism is a multi-dimensional topic, some further studies are needed which utilise multi-disciplinary methodologies and relate the behaviour of absentees to their social origins and educational backgrounds.

Unfortunately, too little is generally known about the processes which lead to a decline and disintegration in the conduct of certain pupils—including absentees and truants—despite some recent studies which will help teachers to overcome the sort of situations they can and do meet in their classrooms. This work will not be easy given the methodological difficulties involved in obtaining reliable information (Wright, 1973). However, if such work is not undertaken, it may not be too long before the notorious comments of one judge that each teacher should expect to be assaulted at least once in his career are fulfilled. An unhappy prospect!

REFERENCES

Billington, B. J. (1979). Truants: some personality characteristics. *Durham and Newcastle Res. Rev.*, 9, 43, 1-6.

Bird, C., Chessum, R., Furlong, J., and Johnson, D. (Ed.) (1980). *Disaffected Pupils.* Brunel: Brunel University Educational Studies Unit.

Buist, M. (1980). Truants talking. *Scot. educ. Rev.*, 12, 40-51.

Brooks, E. E., Buri, J., Byrne, E. A., and Hudson, M. C. (1962). Socioeconomic factors, parental attitudes and school attendance. *Social Work*, 7, 103-108.

Carroll, H. C. M. (1977). Absenteeism in South Wales: studies of pupils, their homes and their secondary schools. *Univ. College, Swansea, Faculty of Education J.*

Caspari, I. (1976). *Troublesome Children in the Class.* London: Routledge and Kegan Paul.

Davie, R. (1980). Promoting school adjustment. In Pringle, M. K. (Ed.), *A Fairer Future for Children.* London: Macmillan.

Davie, R., Butler, N., and Goldstein, H. (1972). *From Birth to Seven.* London: Longmans in association with National Children's Bureau.

D.E.S. (1982). *The New Teacher in School.* A report by HMI. London: HMSO.

Dunham, J. (1981). Disruptive pupils and teacher stress. *Educ. Res.*, 23, 205-213.

Educational Institute of Scotland (1975). Truancy and indiscipline. *Scot. educ. J.*, 7, 256-258.

Elliott, R. (1975). Some characteristics of school non-attenders at Lisnevin School. *Community Home School Gazette*, 69, 401-403.

Evans, M. (1981). *Disruptive Pupils.* London: Schools Council.

Ferri, E. (1976). *Growing-up in a One-Parent Family.* Windsor: N.F.E.R.

Fogelman, K., and Richardson, K. (1974). School attendance: some results from the National Child Development study. In Turner, B. (Ed.), *Truancy.* London: Ward Lock Educational.

Galloway, D. (1976). Size of school, socio-economic hardship, suspension rates and persistent unjustified absence from school. *Br. J. educ. Psychol.*, 46, 40-47.

Galloway, D., Seyd, R., and Ball, T. (1978). Absence from school and behaviour problems in school. *Therapeutic Educ.*, 6, 18-34.

Galloway, D., Ball, T., Blomfield, D., and Seyd, R. (1982). *Schools and Disruptive Pupils.* London: Longmans.

Grunsell, D. (1978). *Born to be Invisible: The Story of a School for Truants.* London: Macmillan Education.

Hargreaves, D., Hestor, S., and Mellor, F. (1975). *Deviance in Classrooms.* London: Routledge and Kegan Paul.

Hersov, L. (1960). Persistent non-attendance at school. *J. child Psychol. Psychiat.*, 1, 130-136.

H.M.I. (1978). *Truancy and Behavioural Problems in Some Urban Schools.* London: Department of Education and Science.

I.L.E.A. (1981). *Perspective on Attendance.* Inner London Education Authority Research and Statistics Branch Document RS/749/80, London.

Johnson, D., Ransom, E., Packwood, T., Bowden, K., and Kogan, M. (1980). *Secondary Schools and the Welfare Network.* London: Unwin Educational.

Jones, A. (1980). The school's view of non-attendance. In Hersov, L., and Berg, I. (Eds.), *Out of School.* Chichester: Wiley.

Jones-Davies, C., and Cave, R. (Eds.) (1976). *The Disruptive Pupil in the Secondary School.* London: Ward Lock Educational.

Kelly, G. (1955). *The Psychology of Personal Constructs* (2 vols.). New York: Norton.

Laslett, R. (1977). Disruptive and violent pupils: the facts and the fallacies. *Educ. Rev.*, 29, 152-162.

Lawrence, J., Steed, D., and Young, P. (1977). *Disruptive Behaviour in a Secondary School.* Educational Studies Monograph, Goldsmith College, London.

May, D. (1975). Truancy, school absenteeism and delinquency. *Scot. educ. Stud.*, 7, 97-107.

N.A.S/U.W.T. (1981). *Discipline or Disorder in Schools: a Disturbing Choice.* N.A.S/U.W.T.

N.A.T. (1975). *Report on Truancy and Disruptive Behaviour in Schools.* National Association of the Teachers of Wales.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., and Bent, D. H. (1975). *SPSS—Statistical Package for the Social Sciences.* 2nd edition. New York: McGraw-Hill.

Pack Report (1977). *Truancy and Indiscipline in Secondary Schools.* Edinburgh: HMSO.

Patrick, H., Bernbaum, G., and Reid, K. (1983). *The Structure and Process of Initial Teacher Education in Universities in England and Wales.* A report on a DES funded project. Leicester: University of Leicester School of Education.

Reid, K. (1981). Alienation and persistent school absenteeism. *Res. in Educ.*., 26, 31-40.

Reid, K. (1982a). The self-concept and persistent school absenteeism. *Br. J. educ. Psychol.*, 52, 179-187.

Reid, K. (1982b). School organisation and persistent school absenteeism: an introduction to a complex problem. *School Organisation*, 2, 45-52.

Reid, K. (1982c). Persistent school absenteeism. *Westminster Studies in Educ.*, 5, 27-36.

Reid, K. (1983a). Institutional factors and persistent school absenteeism. *Educ. Manage. Admin.*, 11, 17-27.

Reid, K. (1983b). Retrospection and persistent school absenteeism. *Educ. Res.*, 25, 110-115.

Reid, K. (1983c). Disruptive persistent absentees. In Frude, N. and Gault, H. (Eds.), *Children's Aggression from School.* Chichester: Wiley.

Reid, K. (1983d). Differences between the perceptions of persistent absentees towards parents and teachers. *Educ. Stud.*, 9, 211-219.

Reid, K., and Kendall, L. (1982). A review of some recent research into persistent school absenteeism. *Br. J. educ. Stud.*, XXX, 295-314.

Reid, K., and Patrick, H. (1980). The structure and process of initial teacher education in universities. In Alexander, R., and Whittaker, J. (Eds.), *Recent Developments in PGCE Courses.* London: SRHE.

Reid, K., Bernbaum, G., and Patrick, H. (1981). On course: students and the PGCE. Unpublished paper presented at the UCET conference, Oxford, November.

Reynolds, D., Jones, D., and St Leger, S. (1976). Schools do make a difference, *New Society*, 37, 721, 223-225.

Reynolds, D. (1982). A state of ignorance? *Educ. for Devel.*, 7, 2, 4-35.

Rutter, M. (1967). A children's behaviour questionnaire for completion by teachers: preliminary findings. *J. child Psychol. Psychiat.*, 8, 1-11.

Rutter, M., Ouston, J., Maughan, B., and Mortimore, P. (1979). *15,000 Hours.* London: Open Books.

Stott, D. H. (1966). *Studies of Troublesome Children.* London: Tavistock.

Stubbs, M., and Delamont, S. (Eds.) (1976). *Explorations in Classroom Observation.* London: Wiley.

Tattum, D. (1982). *Disruptive Pupils in Schools and Units.* Chichester: Wiley.

Tennent, T. G. (1971). School non-attendance and delinquency. *Educ. Res.*, 13, 185-190.

Tibbenham, A. (1977). Housing and truancy. *New Society*, 39, 753, 501-502.

Trigg, J. E. (1973). Focus on the absent minority. *Spec. Educ.*, 62, 2, 24-28.

Tyerman, M. J. (1968). *Truancy.* London: University of London Press.

U.C.A.C. (1975). *A Report by a Sub-committee of Members of U.C.A.C. on Truancy and Disruptive Behaviour in Schools.* Cardiff: The Society.

Watkin, A. (1975). Truancy in Wales: facts and figures. *Education*, 146, 423-424.

West, D. (1982). *Delinquency: Its Roots, Careers and Prospects.* London: Heinemann Educational.

Wright, D. (1973). The punishment of children. In Turner, B. (Ed.), *Discipline.* London: Ward Lock Educational.

(Manuscript received 6th February, 1984)

Br. J. educ. Psychol., 54, 331-335, 1984

A FACTOR ANALYTIC STUDY OF THE McCARTHY SCALES OF CHILDREN'S ABILITIES

By M. TRUEMAN, A. LYNCH, AND A. BRANTHWAITE
(*Department of Psychology, University of Keele*)

SUMMARY. The anglicised version of the McCarthy Scales of Children's Abilities (MSCA) was administered to 89 4¼-year-old children in Stoke-on-Trent. Factor analysis gave some support for the construct validity of McCarthy's General Cognitive, Perceptual-Performance, Verbal and Motor Indexes. However, no support was found for Memory or Quantitative Indexes. These results are considered in the light of previous studies of the construct validity of the MSCA, and it is suggested that results from the Quantitative Scale, in particular, should be interpreted cautiously with children under the age of 5 years.

INTRODUCTION

THE McCarthy Scales of Children's Abilities (McCarthy, 1972) are individually administered measures of cognitive, spatial, and motor abilities. Designed in the USA for use with 2½- to 8½-year-olds, the MSCA was developed in order to provide diagnostic profiles of ability for educational and clinical remedial programmes (Kaufman and Kaufman, 1977). The MSCA consists of 18 short subtests (two of which are further divided into two parts), weighted and combined to yield five Scaled Indexes. These Scaled Indexes separately evaluate the child's verbal, perceptual, quantitative, memory and motor abilities (see Table 1 for comparison of these indexes). Furthermore, the first three scales are combined to yield a General Cognitive Index which is a measure of overall cognitive functioning similar to an IQ score.

McCarthy used intuitive and functional considerations together with the results of factor analysis to determine the placement of each test item on one or more of the scales (McCarthy, 1972; Kaufman and Hollenbech, 1973). As diagnostic decisions concerning assessment and remediation are made on the basis of the child's ability profile, it is clearly important to establish the construct validity of the individual scales. This is particularly so since research in the UK has revealed a number of differences between the performance of American and English children on the MSCA (Lynch *et al.*, 1982a, 1982b). A suitable empirical test of the MSCA's structure can be attempted using factor analysis, the results of which should reflect the structure shown in Table 1. That is to say there ought to be a General Cognitive Factor (probably in the unrotated solution), and five specific factors following rotation corresponding to the five scales.

TABLE 1

TEST ITEMS WHICH MAKE UP THE FIVE SPECIFIC ABILITY SCALES OF THE MSCA

Verbal	Perceptual-Performance	Quantitative	Memory	Motor
Pictorial Memory	Block Building	Number Questions	Pictorial Memory	Leg Co-ordination
Word Knowledge	Puzzle Solving	Numerical Memory 1+2	Tapping Sequence	Arm Co-ordination
Verbal Memory 1+2	Tapping Sequence	Counting & Sorting	Verbal Memory	Imitative Action
Verbal Fluency	Right-Left Orientation		Numerical Memory	Draw-a-Design
Opposite Analogies	Draw-a-Design			Draw-a-Child
	Draw-a-Child			
	Conceptual Grouping			

N.B. The 15 test items that constitute the first three scales are combined into the General Cognitive Index.

Factor analysis of the total standardisation sample divided into five separate age levels (2½, 3-3½, 4-4½, 5-5½, and 6½-8½) yielded results which were quite consistent with the final test structure (Kaufman, 1975). Readily interpretable factors emerged for General Cognitive, Verbal, Memory and Motor abilities at each age level. A Perceptual-Performance factor was isolated in all but the youngest age group. However, a Quantitative factor was only identified in the two oldest groups. There are also some methodological and interpretative weaknesses in this confirmatory study.

A study by Wiebe and Watkins (1980) raised some doubts about the construct validity of the MSCA. They carried out a factor analysis on data taken from some 200 American pre-schoolers. Their interpretation gave empirical evidence only for a General Cognitive factor, whilst the other three resultant factors could not be readily interpreted by Wiebe and Watkins within the MSCA framework. It would seem useful to carry out further investigation of the MSCA, particularly since the present study is just the third separate large sample on which the McCarthy Scales have been employed. In addition, it is important to test the validity of the MSCA with respondents from different cultures, which has not been done previously.

METHOD

Sample

The study population comprised 45 boys and 44 girls living in North Staffordshire who were born between May, 1973, and July, 1975 (overall mean age ± SD at testing = 50·5 ± 1·4 months). Further details of the sample and test procedure can be found in the paper by Lynch *et al.* (1982a).

Analysis

The data used in this analysis comprised the weighted raw scores for the individual test items following the recommended scoring system (McCarthy, 1972). Part two of the Numerical Memory was omitted because only two children managed to score on this subtest. Finally, right-left orientation was not appropriate for use with this age group.

In order to increase the ratio of subjects to variables, the two parts of the Verbal Memory subtests were combined to yield a single variable score. This seemed theoretically justifiable as the subtests assess very similar abilities in terms of Guilford's Structure of Intellect model (Kaufman, 1973).

Pearson product moment correlations were calculated between the MSCA subtests and the resultant correlation matrix was employed in an iterative principal factor analysis. Squared multiple correlations were used as the communality estimates in the main diagonal. Factors with eigen-values greater than unity were extracted and then rotated according to the varimax criterion. Variables with loadings greater than 0·30 were selected for the purposes of factor interpretation.

RESULTS

Table 2 shows the loadings for the first unrotated factor. This first factor seems to provide evidence for a general factor, as it accounts for almost four times as much co-variance as the second factor. This factor would also appear to correspond with the hypothesised General Cognitive factor because all of the relevant variables load significantly (see Table 2). However, the unexpectedly high loadings for the leg co-ordination and imitative action subtests show that the results are not as clear-cut as might be hoped as these variables are motor tests and have little apparent cognitive

TABLE 2

FACTOR MATRIX FOR THE MSCA

MSCA Subtest	Factor 1 UN	 VX	Factor 2 UN	 VX	Factor 3 UN	 VX	Factor 4 UN	 VX	Communality
Block Building	34	34	17	24	09	−11	−18	06	19
Puzzle Solving	57	57	25	27	−06	08	−09	02	40
Pictorial Memory	50	12	−27	53	03	18	−14	15	35
Word Knowledge	53	16	−24	45	16	18	−01	31	36
Number Questions	42	21	−11	40	−24	23	−15	−13	27
Tapping Sequence	45	43	19	18	08	04	00	16	24
Verbal Memory	67	19	−30	78	20	02	−35	24	70
Leg Co-ordination	50	10	−27	34	45	16	19	64	56
Arm Co-ordination	14	03	−05	−02	48	−03	24	55	31
Imitative Action	32	21	03	14	33	−02	08	39	22
Draw-a-Design	56	77	55	−04	14	02	18	25	66
Draw-a-Child	56	72	45	00	−12	20	18	05	57
Numerical Memory	57	46	13	36	06	03	−13	12	36
Verbal Fluency	48	10	−29	41	00	31	05	20	31
Counting and Sorting	64	48	05	47	−32	26	−19	−18	55
Opposite Analogies	62	14	43	31	−49	94	45	03	100
Conceptual Grouping	67	55	13	31	−16	31	05	05	50
% Common Factor Variance	60·9%		16·5%		14·3%		8·3%		
% Total Variance	45·9%		12·4%		10·8%		6·2%		

UN = Unrotated; VX = Varimax rotated.

content. We can conclude, nonetheless, that there is substantial (but not unequivocal) evidence for the existence of a general factor which is similar to McCarthy's General Cognitive Index.

Following rotation of the factor matrix the first rotated factor is a Perceptual-Performance factor on which all the relevant McCarthy subtests load significantly (see Table 2). This interpretation is not invalidated by the quite high loadings for the 'Numerical Memory' and 'Counting and Sorting' subtests, particularly since the latter subtest requires predominantly non-verbal manipulation of concrete materials for its completion, as do the other Perceptual-Performance subtests.

The second rotated factor is rather more complex. The highest loading variables would suggest a verbal label; however, there are also significant loadings for all of McCarthy's Quantitative and most of the Memory subtests. Nevertheless, this factor is clearly distinguishable from a General Cognitive factor because it has only one Perceptual-Performance subtest loading significantly. This factor was designated a hybrid Verbal/Quantitative factor in spite of its symbolic memory content. This label was applied because it reflects previous research showing that Verbal and Quantitative processes are frequently undifferentiated at this age (Thurstone and Thurstone, 1948; Kaufman, 1975).

The third rotated factor does not directly reflect any of McCarthy's Scales. This factor was labelled a Verbal Reasoning factor because the three significantly loading subtests ('Opposite Analogies', 'Conceptual Grouping', and 'Verbal Fluency') all assess classificatory and relational thinking ability whilst relying heavily on verbal content.

The fourth rotated factor is a Gross Motor factor. Three of McCarthy's five Motor subtests load significantly 'Leg co-ordination', 'Arm co-ordination', and 'Imitative action' and these exclude the two fine motor drawing tasks. We can offer no explanation for the anomalous significant loading for the Word Knowledge subtest on this factor.

DISCUSSION

The present study provides generally good evidence for a general factor broadly similar to the *General Cognitive Index* and also a clear *Perceptual-Performance* factor. Only a hybrid *Verbal/Quantitative* factor was isolated rather than two distinct factors. The highest loading variables support the construct validity of a Verbal Scale. However, the other significantly loading variables indicate that caution is required when evaluating the distinctiveness of numerical ability in young children with the Quantitative Scale. The *Verbal Reasoning* factor isolated here was similar to factors isolated in the oldest of Kaufman's five age groups (6½-8½ years) as well as in the Wiebe and Watkins study.

The *Gross Motor* factor underlines Kaufman's (1975) finding that the McCarthy Motor factors tend to be defined by the leg and arm co-ordination as well as by the imitation subtests, rather than by the fine motor co-ordination drawing tests. Indeed, there is little evidence to date from factor analytic studies to support the placement of the drawing tests on McCarthy's Motor Scale.

One noticeable feature of our results is the absence of a *Memory* factor. This was unexpected as such a factor had been isolated in all of Kaufman's (1975) analyses. However, close examination of his memory factors shows they were often complicated by high loading non-memory variables. As Memory Scale subtests all overlap with other scales, it is perhaps not surprising that clearly defined memory factors are not isolated. However, this finding is also a reminder that memory underlies a variety of mental tasks of a simple verbal or numerical kind (especially for 4-year-olds).

CONCLUSIONS

To obtain a wide ranging description of a child's performance from a single test such as the MSCA is of considerable value to the educational or clinical psychologist (Eiser, 1978). Confidence in the ability-profile derived from the test would be enhanced by factorial support for the Scales. However, despite the findings of Kaufman (1975), the present study and that of Wiebe and Watkins (1980) point to certain limitations with the MSCA.

Combining the results of the present study with the cumulated findings from previous research leads us to conclude that:

(1) there is satisfactory factorial evidence to support the construct validity of the General Cognitive Index, Perceptual-Performance Scale and the Verbal Scale;

(2) the Quantitative Scale should be interpreted cautiously with children under 5 years as it may not be distinct from other scales, particularly the Verbal Scale;

(3) McCarthy's Motor Scale should be regarded as a measure of gross motor ability;

(4) lastly, the Memory Scale may not be independent of other scales, such as verbal and numerical.

Whilst this and previous research indicates that some of these scales may not be factorially independent, this does not preclude us from separating the items into distinct scales for pragmatic reasons, providing that it is borne in mind that the resulting measures of ability are interdependent and rely on competence in other areas to achieve their expression.

ACKNOWLEDGMENTS.—This work was supported by grants to the second author and Professor S. J. Hutt from the Medical Research Council and Action Research for the Crippled Child. The invaluable help towards this study of E. M. Vincent and L. B. Mitchell is also gratefully acknowledged. We are also indebted to James Hartley, John Hutt, and John Hegarty for comments on an earlier draft of this paper.

An extended version of this paper is available on request.

REFERENCES

EISER, C. (1978). An evaluation of the McCarthy Scales of Children's Abilities. *Br. J. educ. Psychol.* 48, 351-353.

KAUFMAN, A. S. (1973). Analysis of the McCarthy Scales in terms of Guilford's structure of intellect model. *Percept. mot. Skills*, 36, 967-976.

KAUFMAN, A. S. (1975). Factor structure of the McCarthy Scales at five age levels between 2½ and 8½. *Educ. psychol. Measur.*, 35, 641-656.

KAUFMAN, A. S., and HOLLENBECH, G. P. (1973). Factor analysis of the standardisation edition of the McCarthy Scales. *J. clin. Psychol.*, 29, 358-362.

KAUFMAN, A. S., and KAUFMAN, N. L. (1977). *Clinical Evaluation of Young Children with the McCarthy Scales*. New York: Grune and Stratton.

LYNCH, A., MITCHELL, L. B., VINCENT, E. M., TRUEMAN, M., and MACDONALD, L. (1982a). The McCarthy Scales of Children's Abilities: a normative study on English four-year-olds. *Br. J. educ. Psychol.*, 52, 133-143.

LYNCH, A., VINCENT, E. M., MITCHELL, L. B., and TRUEMAN, M. (1982b). How 'flat' are normal ability profiles in 4-year-olds? *Child Care, Health and Development*, 8, 39-49.

MCCARTHY, D. (1972). *Manual for the McCarthy Scales of Children's Abilities*. New York: The Psychological Corporation.

THURSTONE, L. L., and THURSTONE, T. G. (1948). *Examiner Manual for the S.R.A. Primary Mental Abilities: Primary*. Chicago: Science Research Associates.

WIEBE, M. J., and WATKINS, E. O. (1980). Factor analysis of the McCarthy Scales of Children's Abilities on preschool children. *J. School Psychol.*, 18, 154-162.

(*Manuscript received 22nd February*, 1984)

RESEARCH NOTE

Br. J. educ. Psychol., 54, 336-339, 1984

DIFFERENCES BETWEEN PREFERRED AND ACTUAL CLASSROOM ENVIRONMENT AS PERCEIVED BY PRIMARY STUDENTS AND TEACHERS

By BARRY J. FRASER
(*Western Australian Institute of Technology*)

Summary. This study of perceptions of classroom psychosocial environment is distinctive in that it involved a sample at the primary school level and it incorporated assessments of both preferred and actual environment. Administration of a new short version of the My Class Inventory to a sample of 758 third grade students and their 22 teachers supported the validity of the instrument and revealed some interesting systematic differences between pupils and teachers in their perceptions of preferred and actual classroom environment. In particular, both pupils and teachers tended to prefer a more favourable classroom environment than was perceived as being actually present, and teachers generally perceived the environment of their classes more positively than did their pupils in the same classrooms.

INTRODUCTION

Several comprehensive reviews (Fraser, 1981; Fraser and Walberg, 1981) summarise the considerable interest over the past 15 years in the conceptualisation, assessment and study of perceptions of psychosocial aspects of classroom learning environments. Past studies which have used pupil perceptions of classroom environment as criterion variables have revealed meaningful differences between classes following different curriculum materials and interesting variations in classroom environments according to class size, subject matter and grade level. But the strongest tradition in past classroom environment research has indicated consistently that pupils' classroom environment perceptions account for appreciable amounts of variance in learning outcomes, often beyond that attributable to pupil characteristics such as pretest performance, general ability or both (Fraser and Fisher, 1982).

A large proportion of prior classroom environment studies have been confined to samples of pupils at the secondary school level and to assessing perceptions of what classrooms are *actually* like. It is only very recently that researchers also have assessed what pupils *prefer* their classrooms to be like. For example, Fraser and Fisher (1983) investigated the person-environment fit hypothesis of whether pupils achieve better in their preferred classroom environment, and teachers used information about discrepancies between actual and preferred classroom environment as a practical basis to guide improvements in their classrooms (Fraser, 1983). The present study aimed to investigate differences between primary school pupils and their teachers in their perceptions of preferred and actual classroom environment. The study is distinctive in that it involved primary school classes and it incorporated both actual and preferred forms of environment scales.

ASSESSING CLASSROOM PSYCHOSOCIAL ENVIRONMENT

The instrument most suited to measuring primary school pupils' perceptions of actual classroom environment is the *My Class Inventory* (*MCI*) (Fisher and Fraser, 1981). It has 38 items measuring five dimensions (namely, Satisfaction, Friction, Competitiveness, Difficulty, and Cohesiveness), has item wording well suited to primary school pupils, and has satisfactory scale reliabilities. A typical item in the Difficulty scale is " Schoolwork is hard to do " and a typical item in the Cohesiveness scale is " Children in our class like each other as friends ". Items are answered with a two-point (Yes-No) response format.

Because some researchers and numerous teachers have found that administration of the original MCI can take up a relatively large amount of valuable lesson time especially among younger children, Fraser (1982a) developed a short 25-item version to provide a more rapid assessment of classroom environment. This was based largely on item analyses performed on data collected with the MCI's long form for a large Australian sample of over 2,000 12-year-olds.

The short form was cross-validated in the present study by administrating both the actual and preferred versions to a sample of 758 Grade 3 pupils in 32 classes in eight schools in suburbs of Sydney, Australia. Because some reading difficulties could have been experienced by pupils of this age, a research assistant visited each school to administer the questionnaire by reading items aloud to the whole class. Pupils in the present research were younger than in almost any previous classroom environment study.

Estimates were made of each scale's internal consistency reliability (alpha coefficient) and discriminant validity (mean magnitude of the correlation of a scale with the other four scales). Because the class mean is normally recommended as the unit of analysis in research involving the short form of the MCI, information is reported for class means. Alpha coefficients for different scales for this sample ranged from 0·58 to 0·81 for the actual form and from 0·60 to 0·82 for the preferred form. Although these values are somewhat smaller than those reported for the long form, they are still satisfactory, especially for scales containing only five items. The short form of each MCI scale also displayed good discriminant validity, with values of the mean magnitude of the correlation of a scale with the other four scales ranging from 0·11 to 0·31 for the actual form and from 0·30 to 0·38 for the preferred form. These values indicate satisfactory discriminant validity and suggest that the short form of the MCI measures distinct, although somewhat overlapping, aspects of classroom environment. Further analyses showed that each MCI actual scale differentiated significantly ($P < 0{\cdot}001$) between classrooms, and that the proportion of variance in MCI scores attributable to class membership ranged from 0·15 to 0·33 for different scales.

METHOD AND RESULTS

An investigation was made of differences between pupil and teacher perceptions of the same actual classroom environment and of discrepancies between preferred and actual environment. This involved a subsample drawn from the 32 third grade classes previously described.

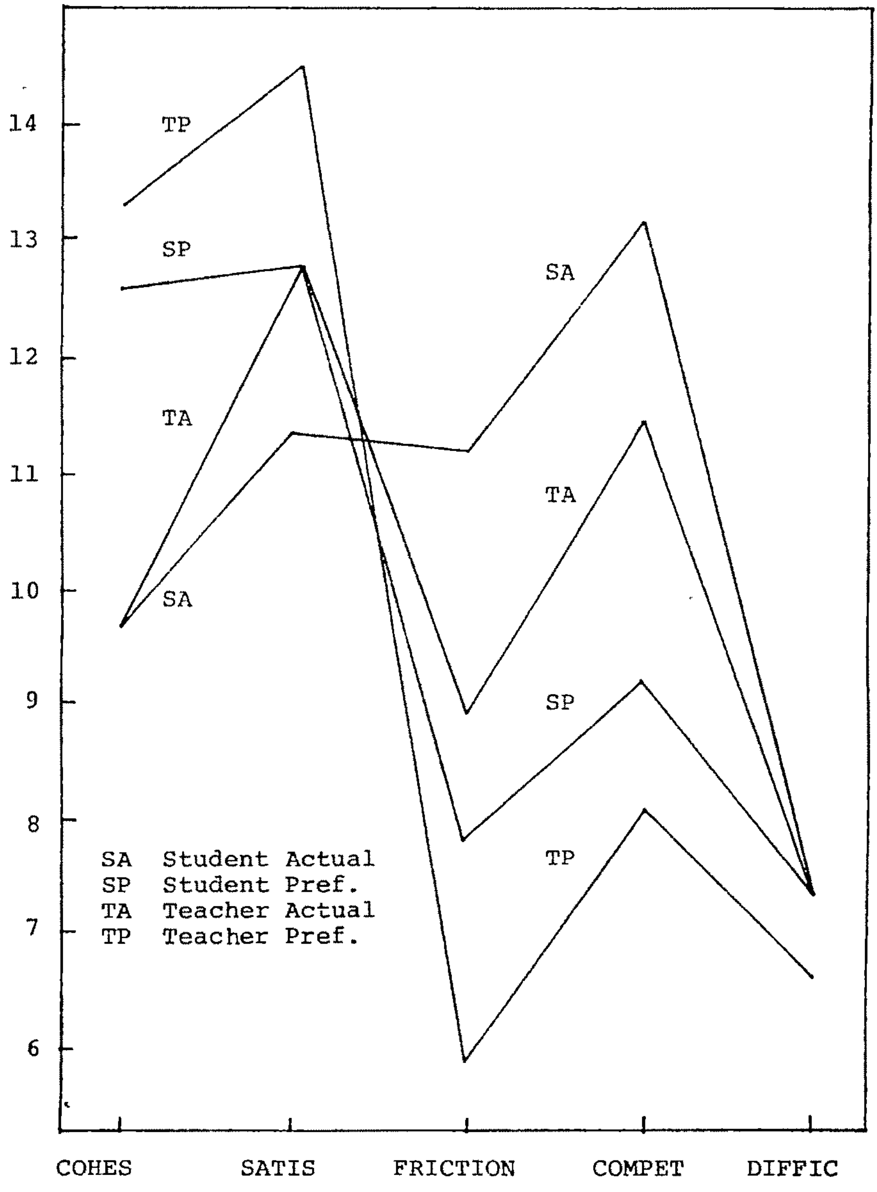
The subsample consisted of 22 classes whose teachers provided responses to both the actual and preferred forms of the MCI. As Figure 1 illustrates, data consisted of 22 teachers' scores on the actual and preferred form of each MCI scale and the 22 pupil class means (both actual and preferred) for the classes taught by these 22 teachers.

Figure 1 shows simplified plots of statistically significant differences between forms. The first stage in the construction of these profiles was the performance of a one-way repeated measures MANOVA for the four-level variable representing the form of the MCI (pupil actual, pupil preferred, teacher actual, teacher preferred). The multivariate test using Wilk's lambda criterion revealed a significant difference ($P < 0{\cdot}01$) among the four forms. To interpret the significant findings for the four-level form factor, a series of t-tests for dependent samples (using the conventional 0·05 level of confidence) was used to test pairwise comparisons between the different forms. (This approach was adopted with the present limited sample size of 22 classes because it seemed to combine the good power characteristics of individual t-tests with the protection against overall Type I error afforded by the requirement that the overall F meet the 0·05 significance criterion; with larger samples, Tukey's or Newman-Keul's tests would be preferred.) In an attempt to provide a more parsimonious picture of differences between forms, only statistically significant differences are plotted in Figure 1. Consequently, any non-significant difference is represented as a zero difference by averaging the relevant pair of scores.

The interpretation of the results in Figure 1 for the comparison of student actual and student preferred forms is that pupils preferred a more favourable classroom environment than they perceived as being actually present on four of the five dimensions assessed by the MCI. In fact, pupils preferred more Cohesiveness, more Satisfaction, less Friction and less Competitiveness. Examination of the profiles of teacher actual and teacher preferred scores indicates that, in comparison with the classroom environment perceived as being actually present, teachers would prefer more Cohesiveness, more Satisfaction, less Friction,

FIGURE 1

PROFILES COMPARING PUPIL AND TEACHER PERCEPTIONS OF ACTUAL AND PREFERRED CLASSROOM ENVIRONMENT



less Competitiveness and less Difficulty. A comparison of pupil actual and teacher actual profiles suggests that teachers perceived a more favourable actual classroom environment in terms of more Satisfaction, less Friction and less Competitiveness than did their pupils in the same classrooms.

CONCLUSION

The use of an economical new short form of the My Class Inventory in comparing pupil and teacher perceptions of actual and preferred classroom environment revealed two interesting patterns of findings. First, in comparison with the emphasis they perceived as being actually present, both pupils and teachers tended to prefer a more positive environment in terms of most of the scales assessed. Second, teachers perceived their classes more favourably on numerous environment scales than did their pupils in the same classrooms. These results replicate prior research in secondary school classrooms (Moos, 1979; Fraser, 1982b) and inform educators that pupils and teachers are likely to differ in the way they perceive the actual environment of the same classrooms, and that the environment preferred by pupils and teachers commonly differs from that actually present in classrooms.

REFERENCES

FISHER, D. L., and FRASER, B. J. (1981). Validity and use of My Class Inventory. *Sci. Educ.*, 65, 145-156.

FRASER, B. J. (1981). *Learning Environment in Curriculum Evaluation: A Review*. Evaluation in Education series. Oxford: Pergamon.

FRASER, B. J. (1982a). Development of short forms of several classroom environment scales. *J. educ. Meas.*, 19, 221-227.

FRASER, B. J. (1982b). Differences between student and teacher perceptions of actual and preferred classroom learning environment. *Educ. Evaluation and Policy Analysis*, 4, 511-519.

FRASER, B. J. (1983). Improving classrooms through the evaluation of learning environments. Paper presented at Annual Meeting of American Educational Research Association, Montreal.

FRASER, B. J., and FISHER, D. L. (1982). Effects of classroom psychosocial environment on student learning. *Br. J. educ. Psychol.*, 52, 374-377.

FRASER, B. J., and FISHER, D. L. (1983). Student achievement as a function of person-environment fit: a regression surface analysis. *Br. J. educ. Psychol.*, 53, 89-99.

FRASER, B. J., and WALBERG, H. J. (1981). Psychosocial learning environment in science classrooms: a review of research. *Studies Sci. Educ.*, 8, 67-92.

MOOS, R. H. (1979). *Evaluating Educational Environments: Procedures, Measures, Findings and Policy Implications*. San Francisco: Jossey-Bass.

(*Manuscript received 6th February*, 1984)

BOOK REVIEWS

BRUNER, J. (1983). *In Search of Mind: Essays in Autobiography.* London: Harper and Row, pp. 306, £14·00.

Recently, the Sloan Foundation have had the idea of supporting a number of distinguished scientists, from various areas, to write their autobiographies. This volume is Jerry Bruner's contribution to the enterprise. As one might expect, it is attractive even from the purely literary point of view. In addition, it provides a brief and comprehensible account of the contribution he himself regards as major in his work, as well as some interesting light on the processes that formed an originator and leader in psychology.

There is a problem in organising such a book, which he has tackled by dividing it into three parts. In the first, he goes chronologically through the events of his life up to the point when he returned from the army to Harvard. In the second part, he devotes a chapter to each of the intellectual areas in which he has worked since then, letting the events of his life come in where they naturally do so. Thus, there is a chapter on selective and biasing factors in perception (the " New Look "), another on reasoning and concept formation (" A Study of Thinking "), one on the development of thinking in the infant (" Stages of Cognitive Growth "), and one on language (" Child's Talk "). Finally, there is one on teaching (" The Process of Education "). As those who know the books will realise, the different intellectual interests did on the whole follow each other, but not totally. Within each chapter therefore he ranges back and forth in time, raising relevant people and events as they become appropriate. Finally, there is a third section in which the whole life is cross-classified in another way, by interests outside psychology, by particular friends, and by universities.

This method has some advantages, but also leaves some intriguing gaps. How far, for instance, did the enormous public success of his educational writings alter his later attitude to research on language? Conversely, did the profitable collaborations with Postman and with George Miller in the 1950s and 1960s create a need to repeat those experiences? Did he therefore wish to tackle later problems by similar means but in a different and perhaps inappropriate context? After the stimulating conversation and the burst of ideas, there comes the need to work out the implications, collecting data and hammering out the precise form of a theory. Many of us would hold that the very success of the cognitive revolution required such a stage. A hint of Bruner's own attitude is to be found in his comment that, when he was deciding to leave Harvard in the early 1970s, " the era of Great Men . . .was over ". Similarly, he was clearly disappointed by the absence in Oxford of the general discussions that had been the glory of the Center for Cognitive Studies.

So far as his contribution goes, we can accept his own estimate, that he has throughout his career stood up for the processes that a child, or an adult, brings to a situation and imposes upon it. He has indeed been the champion of such top-down processes from times when they were much less popular than they now are; beyond the information given, as he puts it. We all owe him an intellectual debt for this emphasis. As a contribution to the psychology of psychologists, it is valuable that he records that he himself was a subject who showed particularly strongly the bad effects, on correctly perceiving a blurred picture, of starting with a wrong hypothesis. Going back more deeply into his infancy, he was blind until after the age of two, so that his own perceptual processes may well have relied particularly strongly on the theory from inside rather than the sensations from outside. If that was what led him to challenge the false view that all perception is bottom-up, it was a nice example of his ill-wind blowing the rest of us good.

Of course, perception is not solely top-down, and there are some places where one feels that he may have been wrong in his reading of situations and people. This is clearest in his account of events at Oxford, which is frankly unrecognisable to somebody who was there at the time. The same problem comes up elsewhere, however, and in ways that sum up his strength and also perhaps his failings. The book is thickly studded with the names of those he has met, from Kennedy and Johnson to Oppenheimer and Gombrich, from McGeorge Bundy to Alan Bullock or Mary Warnock. All, all, are witty, likeable, and of sound judgment; the only name that comes in for execration is that of Adolf Hitler, which is understandable. One suspects that charm and the style are being imposed by the writer on a world that may not always truly possess them. Such uniform admiration is likeable, but blurs issues.

One detail that he does not mention is this. By his presence in Britain at a time when developmental psychology was beginning to expand, he provided a major international figure to encourage the area, to supply a model, and to prove respectability to outside bodies. We owe him a lot. From this book, one can learn what it felt like from the inside.

D. BROADBENT

ELLIS, A. W. (1984). *Reading, Writing and Dyslexia: a Cognitive Analysis*, London: Lawrence Erlbaum Associates, pp. 147, £12·95.

This book is basically an account of the ways in which cognitive psychologists have studied reading and writing. Topics covered include the English writing system, the recognition of words from their auditory and visual features, the comprehension of sentences, and discussions of both acquired and developmental dyslexia.

One of its great strengths is its extreme lucidity, and this is achieved without any sacrifice of scholarship. The appropriate level of technicality has been gauged with particular skill: there is material for psychologists at all levels from beginner to advanced research worker, and even those with no training at all in psychology will seldom be out of their depth. In addition there is a genuineness about the book: the author comes through not as a purveyor of second-hand information but rather as a researcher who has himself wrestled with the problems and is fully aware of possible sources of erroneous inference.

The account of the varieties of acquired dyslexia in Chapter 4 is particularly helpful. With regard to developmental dyslexia (Chapter 8), I fully agree with Ellis when he says that it is not simply " a disorder of short-term memory " (p. 110), since how, in that case could one explain the uncertainties over ' left ' and ' right ' where memory load is minimal? Similarly, although it is well established that *some* dyslexic children show unusual eye movements it seems equally well established that not *all* of them do so. What seems to emerge from the evidence—and I think Ellis would agree—is that there is a group of language disorders (not just *reading* disorders *simpliciter*) which constitute a recognisable pattern or syndrome even though there is considerable variation between individuals. I do not know if this constitutes a ' unitary explanation ' of dyslexia (cf. pp. 107-110), but I wonder if Ellis would have done better to leave the word ' unitary ' alone. When people discuss whether developmental dyslexia is a ' unitary condition ' it is by no means self-evident what is at issue. The parallel with obesity (p. 107) is a helpful one, but the boundaries of what constitutes a ' disease ' (pp. 106-7) and what constitutes a unity are imprecise. Cancer, for example, is a disease but not necessarily a unity, while pregnancy is presumably a unity (since it always arises from the same cause, viz. insemination), but though temporarily incapacitating and normally requiring skilled medical attention it is not normally thought of as a disease! As Ellis in effect points out, concepts which are of value in medicine or education do not *have* to represent either unities or diseases; indeed it is possible to accept the usefulness of the concept of developmental dyslexia without believing that it represents either.

My only serious worry about the book is the way in which expressions such as ' psychological processes ' and ' mental operations ' are bandied about (as, indeed, they are by others) as if no explanation was needed. This worry is highlighted at the start of Chapter 4 where we are told: " The brain is the organ of the mind. It is also an organ of the body, and as such, is susceptible to injury and illness ". Is ' the mind ' (whatever ' it ' may be) not so susceptible? Ellis seems happily unaware that he is walking over a philosophical minefield!

I noticed two minor errors. ' Phoenician(s) is misspelled on pp. 4 and 75 and the explanation (p. 25) ' Greek . . . genus = birth ' is incorrect: the Greek letter upsilon is standardly transcribed as ' y ' not ' u ', and the nearest Greek word is in fact *genos* which does not mean ' birth ' but either ' race ' or ' sort ' (kind).

On the important points, however, this is an excellent book which I recommend strongly.

T. R. MILES

McKinney, J. D., and Feagans, L. (Eds.) (1983). *Current Topics in Learning Disabilities*, Vol 1, London: Ablex, pp. 306, £33·00.

This is the first volume in a projected series on learning disability whose aim is " to provide a synthesis of current knowledge and theory and thereby strengthen the conceptual foundations underlying contemporary practice ". Each volume, we are told, will comprise three sections. Basic processes will provide critical reviews of research on theoretical advances; contemporary issues (a strange sub-heading in a book entitled 'Current Topics') will consider policy issues and cross-cultural perspectives; perspectives on intervention will describe and evaluate innovative approaches to treatment and service delivery.

In the first section Torgesen and Licht provide a well balanced account of the importance of maladaptive processing behaviour in the field of learning disability. They integrate concepts such as ' the inactive learner ', ' adaptive learning strategies ' and ' metacognition ' in a conceptual framework for understanding the difficulties of *some* LD children on some tasks, thereby avoiding any overstatement of the possible pay-off of this line of research. Feagans adopts a similar stance in her consideration of discourse processes in LD children arguing that, contrary to the traditional isolation of discrete aspects of language, it is language beyond the sentence level which is crucial for school achievement. The review is bedevilled to some extent by a lack of appropriate material to review, such that the many differences found between normal and LD children cannot yet be related to academic achievement or indeed, generalised to all LD.

It is doubtful that the other contributions in this section are ' basic processes ' in the same sense, but this is a criticism of an inadequate classification of contributions rather than a comment on the intrinsic value of the papers. Keogh's ' Individual differences in temperament ' is concerned with the goodness of fit between child and situation and is probably better viewed as an account of the relevance of personality in education generally, than as saying something basic about LD children. Similarly the consideration of LD children's social relationships with significant others (Bryan and Bryan) is interesting but peripheral.

The central section of Contemporary Issues has a less integrated quality. Minimum competence testing (Hall and Gallagher) is largely about policy and the paper on prevention (Karnes and Stoneburner), in selecting general intelligence as the best predictor variable, is clearly not about the same LD as the other chapters. Elkins' chapter is an historical account of the field in Australia, highlighting appropriately, given the general tenor of this volume, the shift in emphasis from learning disability to learning difficulty. It provides an excellent conceptual analysis independent, in fact, of the Australian connection. The theoretical stance of chapters 1 and 4 are reflected in the third section on Intervention. Hallahan, Kneedler and Lloyd provide a general review of self instruction and self-monitoring techniques, and these are relevant to the LD in so far as these latter fail to adopt appropriate strategies. A few studies do point to a general efficacy of the techniques with LD. The other studies in the section are concerned more with specific experimental programmes. Deshler *et al.* are concerned with the teaching of specific and general strategies to LD adolescents while Blank, Berlin and Rose focus on the non-productive patterns of dialogue between adults and children.

The editors and publishers are to be commended for the lay-out and presentation of this volume, if not for the price. The individual contributions too are, without exception, well researched, readable and worthy of inclusion somewhere in the projected series. One must, however, wonder at the balance of the contributions. The preface comments on ' the substantial controversy about the nature of learning disabilities ' and ' the synthesis of knowledge ' yet the volume is almost entirely cognitive-functional with, apart from Elkins, only passing reference to alternative standpoints. Chapter 1, for example is presented in the spirit of complement—but to what? If the series is intended to be a comprehensive account of the area, one might have looked for wider context-setting representation in Volume 1.

H. B. Pont

POTTON, A. (1984). *Screening*. London: Macmillan, pp. viii+62. n.p.

This is one of a series entitled 'Guides to Assessment in Education' under the general editorship of Jack Wrigley. In a well-illustrated and very clearly expressed fashion, Potton introduces the need for screening, the purpose of screening devices and the influence of the medical model on the educational enterprise. The exposition of basic concepts is first-rate and can be strongly recommended for novice students. The analysis of the utility—or as it turns out, the non-utility—of screening devices is an excellent illustration of rigorous thinking. But Potton, whilst dismissive of screening devices, recognises that the problem of early identification of learning difficulties has not gone away. How do we meet it? Potton recommends that what teachers need is vigilance. We will all drink to that of course. But how, specifically, is vigilance to be deployed in this instance? Here Potton hands over the reader to an appendix written by Joan Dean. Dean spins out a familiar fabric of tips for teachers. She presents a woolly garb of indicators of promise and handicap. She has to be woolly. If she were rigorous, one suspects, Potton would get her.

And here this little text becomes a parable of educational knowledge. So little of it bears serious examination. Yet it has to be used in the day to day business of schooling. Maybe Potton and Dean should work more closely together to submit professional knowledge to analysis. In the present book Potton avoids the action, Dean begs all the questions.

C. DESFORGES

WOLFENDALE, S. (1983). *Parental Participation in Children's Development and Education*. New York: Gordon and Breach, $38·50.

This is a well-documented survey which usefully summarises many of the piecemeal innovations and developments since the Plowden Report, assembling much scattered information for the first time. The title gives a fair indication of the contents. Parental participation is defined as an umbrella term to describe all the models and types of liaison between parents, schools and other community institutions that provide for children. As one might expect the projects discussed are virtually all from pre-school and infant/first school stages, with an emphasis on school-based schemes and services. Although Sheila Wolfendale is unable to cite many specific examples of co-operation between the various education and welfare services, she does note that this is an area of current expansion. If one is to adopt a holistic approach to education and family support (and the book rightly advocates this), then positive co-ordination of professional effort is surely essential. Barriers between specialists and parents ought not to be aggravated by barriers between the specialists themselves.

A welcome feature of the book is the way it urges professional workers in the field to cease regarding families as if they were inadequate and passive clients in need of expert direction. Without some kind of 'reciprocity', in which professionals and parents are mutually involved, the latter can easily be made to feel powerless, under-valued and incompetent. All the accumulating evidence now suggests that parents are vitally concerned about their children's education and want to take an active part. A good example reported by Sheila Wolfendale is the burgeoning of projects in which parents help teach their children to read. The very success of such schemes (with children of widely differing abilities and coming from various ethnic and social groups) serves to demonstrate both parental enthusiasm and the potential of this largely untapped teaching force.

However, what the book fails to emphasise sufficiently with regard to some of the projects is the critical and subtle nature of the interaction between the main participants. Creating a feeling of equality and partnership depends not only on shared objectives and activities but, even more crucially, on establishing a sensitive relationship based on empathy and mutual trust.

It is perhaps inevitable that short synopses of many different projects appear on occasion to distort or diminish them so that their relative mertis or scale cannot be judged. Thus the significance of the Pre-school Playgroups Association—with its more than 21 years of experience, and now serving over half the population of 2½-5 year-olds—is not fully brought out. A high percentage of playgroups are after all run by parent committees operating on democratic, family-centred principles. In fact, many of PPA's objectives coincide with those identified in the book: a supporting network for young mothers, opportunities for community action, involvement in decision making, encouragement of parents as educators of their own children, and parental education in a wider more subtle sense.

The author's organisation of a diverse mass of information is clear throughout, though use of the book as a reference source is unfortunately hampered by an inadequate index. The reader might also be a little put off by a good sprinkling of misprints, the miniaturised type-face used for quotations, and a somewhat dry style of writing. Nevertheless this is a valuable quarry of ideas and strategies for professionals and parents alike.

M. FAWCETT

PUBLICATIONS RECEIVED

Books for review have been selected from the following
('n.p.' means 'no price mentioned').

Académie des Sciences Sociales et Politiques de la République Socialiste de Roumanie (Janvier—Juin 1983). *Revue Romaine Des Sciences Sociales.* Serie De Psychologie. Tome 27 No. 1.

Académie des Sciences Sociales et Politiques de la République Socialiste de Roumanie (Juillet—Decembre 1983). *Revue Romaine des Sciences Sociales.* Series de Psychologie. Tome 27, No. 2.

Altman, I., Lawton, M., and Wohlwill, J. F. (Eds.) (1984). *Elderly People and the Environment Advances in Theory and Research (Vol 7).* New York: Plenum, £39·50.

Anderson, S. B., and Helmick, J. S. (Eds.) (1984). *On Educational Testing.* London: Jossey-Bass, £19·95.

Australian Council for educational Research. *Fifty-third Annual Report 1982-83.* Victoria: ACER.

Batten, M. (1983). *Issues of the Eighties: Principals' Perspectives and School Practices.* Victoria: ACER, $7·50.

Beach, R., and Bridwell, L. S. (Eds.) (1984). *New Directions in Composition Research.* London: Guilford Press, £30·95.

Beattie, G. (1983). *Talk: An Analysis of Speech and Non-Verbal Behaviour in Conversation.* Milton Keynes: Open University Press, £8·95.

Bennett, N., Desforges, C., Cockburn, A., and Wilkinson, B. (1984). *The Quality of Pupil Learning Experiences.* London: Lawrence Erlbaum Associates, £9·95 (h/b), £4·95 (p/b).

Bowers, T. (1984). *Management in the Special School.* Beckenham: Croom Helm, £9·95.

Braithwaite, J. (1984). *Explorations in Early Childhood Education: The Mount Druitt Early Childhood Project.* Victoria: ACER, $20·00.

Broadfoot, P. (1984). *Selection Certification and Control: Social Issues in Educational Assessment.* London: Taylor and Francis, £12·95 (cloth), £7·50 (paper).

Bruner, J. (1983). *In Search of Mind: Essays in Autobiography.* London: Harper and Row, £14·00.

Burgess, R. G. (1984). *The Research Process in Educational Settings: Ten Case Studies.* London: Taylor and Francis, £14·00 (cloth), £7·95 (paper).

Cashmore, E. E. (1984). *No Future, Youth and Society.* London: Heinemann Educational Books, £12·95 (cased), £3·95 (paper).

Cell, E. (1984). *Learning to Learn from Experience.* New York: Suny, $34·50 (cloth), $10·95 (paper).

Cullingford, C. (1984). *Children and Television.* Aldershot: Gower, £13·50.

Curran, H. V. (1984). *Nigerian Children: Development Perspectives.* Henley-on-Thames: Routledge and Kegan Paul, £18·95.

Davis, D. R. (1984). *An Introduction to Psychopathology.* Oxford: Oxford University Press, £12·50 (h/b), £6·95 (p/b).

Dillon, R. F., and Schmeck, R. R. (1983). *Individual Cognition: Vol. 1.* New York: Academic Press, $34·00.

Doise, W., and Palmonari, A. (1984). *Social Interaction in Individual Development.* Cambridge: Cambridge University Press, £27·50.

Ellis, A. W. (1984). *Reading, Writing, and Dyslexia: A Cognitive Analysis.* London: Lawrence Erlbaum Associates, £12·95.

Fagin, L., and Little, M. (1984). *The Forsaken Families.* Harmondsworth: Penguin Books, £3·95.

Fletcher, R. (1984). *Education in Society: the Promethean Fire: A New Essay in the Sociology of Education.* Harmondsworth: Penguin Books, £5·95.

Gregory, R. L. (1984). *Mind in Science.* Harmondsworth: Penguin Books, £7·95.

Hooper, F. H., Hooper, J. O., and Colbert, K. K. (1984). *Personality and Memory Correlates of Intellectual Functioning: Young Adulthood to Old Age.* Karger, $28·75.

Kaplan, E., Goodlass, H., and Weintraub, S. (1983). *The Assessment of Aphasia and Related Disorders Complete Package.* Beckenham: Quest Publishing Agency, $30·25.

Kaye, K. (1982). *The Mental and Social Life of Babies.* Brighton: Harvester Press, £22·50.

Laslett, R., and Smith, C. (1984). *Effective Classroom Management.* Beckenham: Croom Helm, £11·95 (h/b), £5·95 (p/b).

Lindsay, G. (1984). *Screening for Children with Special Needs.* Beckenham: Croom Helm, £14·95.

Loftus, G. R., and Loftus, E. F. (1984). *Mind at Play,* New York: Basic Books, £9·95.

Lueptow, L. B. (1984). *Adolescent Sex Roles and Social Change.* New York: Columbia University Press, $45·50.

Mandl, H., Stein, N. L., and Trabasso, T. (Eds.) (1984). *Learning and Comprehension of Text.* Hillsdale, N.J.: Lawrence Erlbaum Associates, £29·95.

McClintock, A. (1984). *Drama for Mentally Handicapped Children.* London: Souvenir Press, £5·95 (p/b), £8·95 (h/b).

McDermott, R. A. (1984). *The Essential Steiner: Basic Writings of Rudolf Steiner.* San Francisco: Harper and Row, $10·95.

McKinney, J. D., and Feagans, L. (eds.) (1983). *Current Topics in Learning Disabilities.* London: Ablex, £33·00.

Millard, D. (1984). *Daily Living with a Handicapped Child.* Beckenham: Croom Helm, £8·95.

Miller, N. (1984). *Bilingualism and Language Disability.* Beckenham: Croom Helm, £14·95.

Mortimore, P. *et al.* (1984). *Behaviour Problems in Schools.* Beckenham: Croom Helm, £7·95.

Nisbet, J., and Watt, J. (1984). *Educational Disadvantage: Ten Years On.* Edinburgh: HMSO, £4·60.

Osborn, A. F., Butler, N. R. and Morris, A. C. (1984). *The Social Life of Britain's Five-Year-Olds. A report of the Child Health and Education Study.* Henley-On-Thames: Routledge Education Books, £19·95.

Painter, F. (1984). *Living with a Gifted Child.* London: Souvenir Press, £8·95.

Perry, D. G., and Bussey, K. (1984). *Social Development.* Hemel Hempstead: Prentice Hall International, £25·60.

Popkewitz, T. S. (1984). *Paradigm and Ideology in Educational Research: The Social Function of the Intellectual.* London: Taylor and Francis, £14·25 (cloth), £7·25 (paper).

Quigley, S. P., and Paul, P. V. (1984). *Language and Deafness.* Beckenham: Croom Helm, £19·95.

Raffe, D. (Ed.) (1984). *Fourteen to Eighteen: The Changing Pattern of Schooling in Scotland.* Aberdeen: Aberdeen University Press, £7·90.

Reed, M. (1984). *Educating Hearing-Impaired Children.* Milton Keynes: Open University Press, £5·95.

Reigeluth, C. M. (1983). *Instructional-Design Theories and Models: An Overview of their Current Status.* London: Lawrence Erlbaum Associates, £22·95.

Reynolds, C. R., and Clark, J. H. (1983). *Assessment and Programming for Young Children with Low-Incidence Handicaps.* New York: Plenum, $42·50.

Rostron, A., and Sewell, D. (1983). *Microtechnology in Special Education.* Beckenham: Croom Helm, £14·95.

Spaulding, W. D., and Coel, J. K. (1984). *Nebraska Symposium on Motivation 1983 Theories of Schizophrenia and Psychosis.* London: University of Nebraska Press, £23·70 (cloth), £14·20 (paper).

Steinberg, E. R. (1984). *Teaching Computers to Teach.* London: Lawrence Erlbaum Associates.

Stephenson, G. M., and Davis, J. H. (1984). *Progress in Applied Social Psychology, Vol 2.* Bognor Regis: Wiley and Sons, £32·95.

Sternber, R. J. (Ed.) (1984). *Advances in the Psychology of Human Intelligence.* Hillsdale, N.J.: Lawrence Erlbaum Associates, £29·95.

Stones, E., (1984). *Supervision in Teaching Education: a Counselling and Pedagogical Approach.* London: Methuen, £11·50 (h/b), £4·95 (p/b).

Sutherland, G. (1984). *Ability, Merit and Measurement Mental Testing and English Education 1880-1940.* Oxford: Oxford University Press, £25·00.

Swash, M., and Schwartz, M. S. (1984). *Biopsy Pathology of Muscle.* London: Chapman and Hall Medical.

Thatcher, J. (1984). *Teaching Reading to Mentally Handicapped Children.* Beckenham: Croom Helm, £7·95.

Trower, P. (1984). *Radical Approaches to Social Skill Training.* Beckenham: Croom Helm, £16·95.

TURNER, J. (1984). *Cognitive Development and Education.* London: Methuen, £2·95.
TYLER, L. E. (1983). *Thinking Creatively.* London: Jossey-Bass, £16·95.
WARREN, B. (1984). *Using the Creative Arts in Therapy.* Beckenham: Croom Helm, £8·95.
WOOLFOLK, A. E., and MCCUNE-NICOLICH (1984). *Educational Psychology for Teachers* (2nd edition). Hemel Hempstead: Prentice Hall International, £22·75.

AUTHOR INDEX TO VOLUME 54, 1984

Axworthy, D.: see Tierney, I.

Branthwaite, A.: see Trueman, M.

Bar-El, Z.: see Zakay, D.

Bar-Tal, D.: Causes of Success and Failure and their Dimensions as a Function of SES and Gender: a Phenomenological Analysis 51

Brown, G. A.: Toward a Typology of Lecturing Styles 93

Bakhtar, M.: see Brown, G. A.

Calev, A.: see Erwin, P. G.

Chapman, J. J.: Academic Self-Concept: Some Developmental and Emotional Correlates in Nine-Year-Old Children 284

Chapman, J. W.: Consistency of Causal Attributions for Expected and Actual Examination Outcome: a Study of the Expectancy Confirmation and Egotism Models 177

Child, D.: see O'Neil, M. J.

Cooper, C.: see Kline, P.

Crispin, L.: The Relevance of Visual Sequential Memory to Reading 24

Cunningham, C. C.: The Relationship between Maternal Ratings of First Word Vocabulary and Reynell Language Scores 160

Downing, J.: When is a Silent Letter Silent? 111

Entwistle, N. J.: see Kozéki, B.

Erwin, P. G.: The Influence of Christian Name Stereotypes on the Marking of Children's Essays 223

Francis, H.: Children's Knowledge of Orthography in Learning to Read 8

Fraser, B. J.: Differences between Preferred and Actual Classroom Environment as Perceived by Primary Students and Teachers 336

French, J. R. W.: see Wood, H. A.

Furnham, A.: Getting a Job: School-leavers' Perceptions of Employment Prospects 293

Goldberg, M.: see Bar-Tal, D.

Hacker, R. G.: A Hierarchy of Intellectual Development in Science 137

Hamilton, W.: see Crispin, L.

Hargie, O. D. W.: Training Teachers in Counselling Skills: the Effects of Microcounselling 214

Harrop, A.: Modifying ' Creative Writing ' in the Classroom 62

Howarth, P. S.: see Wood, H. A.

Ives, S. W.: The Development of Expressivity in Drawing 152

Kingsmill, M. C.: see Wood, H. A.

Kline, P.: The Factor Structure of the Comprehensive Ability Battery 106

Knaani, A.: see Bar-Tal, D.

Kozéki, B.: Identifying Dimensions of School Motivation in Britain and Hungary 306

Kreitler, S.: see Zakay, D.

LaRocque, L.: see Downing, J.

Lawes, M. M.: see Chapman, J. W. (second entry)

Lynch, A.: see Trueman, M.

McCann, C.: see Harrop, A.

McIntyre, T. C.: The Relationship between Locus of Control and Teacher Burnout 235

Marsh, H. W.: Experimental Manipulations of University Student Motivation and their Effects on Examination Performance 206

Michie, S.: Number Understanding in Preschool Children 245

Nash, C.: Identity and Equivalence Conservation: Longitudinal Field Studies in Sequence and Significance 1

Nash, C.: Kindergarden Task Management and Grade 2 Perceptions of Teacher-Learner Roles 221

Oakhill, J.: Inferential and Memory Skills in Children's Comprehension of Stories ... 31

O'Neil, M. J.: Biggs' SPQ: a British Study of its Internal Structure 228

Phillips, C. J.: Use of Abbreviated Cognitive Tests in the Assessment of Children's Special Educational Needs 168

Ratcliffe, S. G.: see Tierney, I.

Reid, K.: The Behaviour of Persistent School Absentees 320

Dos Santos Veiga, J.: see Seddon, G. M.

Saunders, R. J.: The Effect of Pictures on the Aquisition of a Small Vocabulary of Similar Sight-Words 265

Schenk, S. M.: see Van Rossum, E. J.

Seddon, G. M.: The Transferability of Two Pictorial Scientific Tasks between Different Spatial Dimensions 276

Silva, P. A.: see Chapman J. W.

Sjöberg, L.: Interests, Effort, Achievement and Vocational Preference 189

Sloper, P.: see Cunningham, C. C.

Smilansky, J.: External and Internal Correlates of Teachers' Satisfaction and Willingness to Report Stress 84

Smith, L.: see Tierney, I.

Solman, R. T.: see Saunders, R. J.

Tariq, R. H.: see Seddon, G. M.

Tierney, I.: The McCarthy Scales of Children's Abilities—Sex and Handedness Effects in 128 Scottish Five-Year-Olds 101

Timko, H.: see Downing, J.

Trickey, G.: see Crispin, L.

Trueman, M.: A Factor Analytic Study of the McCarthy Scales of Children's Abilities ... 331

Van Rossum, E. J.: The Relationship between Learning Conception, Study Strategy and Learning Outcome 73

Williams, S. M.: see Chapman, J. W. (first entry)

Wood, D. J.: see Wood, H. A.

Wood, H. A.: The Mathematical Achievments of Deaf Children from Different Educational Environments 254

Youngman, M. B.: see Brown, G. A.

Zakay, D.: Cognitive Orientation and Changing the Impulsivity of Children 40

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

British Journal of Psychology
British Journal of Medical Psychology
Journal of Occupational Psychology
British Journal of Social Psychology
British Journal of Clinical Psychology
British Journal of Developmental Psychology
British Journal of Mathematical and Statistical Psychology
The British Journal of Educational Psychology

The *British Journal of Psychology*, the *British Journal of Medical Psychology*, *Journal of Occupational Psychology*, the *British Journal of Developmental Psychology*, the *British Journal of Social Psychology* and the *British Journal of Clinical Psychology* appear in quarterly parts. The 1984 subscription prices, payable in advance, are £56·00 (US$110·00) for the *British Journal of Psychology*; £36·00 (US$64·50) for the *British Journal of Social Psychology*; £35·00 (US$64·50) for the *British Journal of Clinical Psychology*; £33·00 (US$64·50) for the *British Journal of Developmental Psychology*; £46·00 (US$90·50) for the *British Journal of Medical Psychology*; and £39·00 (US$79·50) for the *Journal of Occupational Psychology*.

The *British Journal of Mathematical and Statistical Psychology* is issued twice a year, in May and November. The subscription rate, payable in advance, is £39·00 per volume post free (US$79·50 in the USA).

Subscriptions to any of these journals may be sent to a bookseller or subscription agent, or to The British Psychological Society, The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK. A special price of £257·00 (US$500·00) for all seven journals going to the same address is offered for 1984.

The British Journal of Educational Psychology is issued three times a year, in February, June and November. The prepaid subscription is £15·00 a year, post free, or £5·50 each number. The prepaid subscription price in the USA is $40·00 post free. Subscriptions should be sent to the Scottish Academic Press, 33 Montgomery Street, Edinburgh EH7 5JX.

MANUSCRIPTS

For the *British Journal of Psychology* should be sent to Professor D. E. Blackman, Department of Psychology, University College Cardiff, PO Box 78, Cardiff CF1 1XL.

For the *British Journal of Medical Psychology* should be sent to Dr Sidney Crown, Department of Psychiatry, The London Hospital, Whitechapel, E1 1BB.

For the *Journal of Occupational Psychology* should be sent to Professor P. Warr, MRC/SSRC Social and Applied Psychology Unit, Department of Psychology, The University, Sheffield S10 2TN.

For the *British Journal of Social Psychology* should be sent to Dr A. S. R. Manstead, Department of Psychology, The University, Manchester M13 9PL.

For the *British Journal of Clinical Psychology* should be sent to Dr D. Shapiro, MRC/SSRC Social and Applied Psychology Unit, Department of Psychology, The University, Sheffield S10 2TN.

For the *British Journal of Mathematical and Statistical Psychology* should be sent to Dr C. Leach, Department of Psychology, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU.

For the *British Journal of Developmental Psychology* should be sent to Professor P. E. Bryant, Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD.

For *The British Journal of Educational Psychology* should be sent in duplicate to Professor Geoffrey Brown, School of Education, University of East Anglia, Norwich NR4 7TJ.

Contributors are advised to consult *Suggestions to Contributors* issued by the Standing Committee on Publications of the British Psychological Society and obtainable, price £2·00 post free, from The British Psychological Society, St. Andrew's House, 48 Princess Road East, Leicester LE1 7DR.

All manuscripts should be submitted in duplicate.

OFFPRINTS

Contributors on application receive fifty copies of their papers free. Additional copies may be obtained in units of fifty copies.

REPRINTS

The *Journal* is available in reprint from Swets and Zeitlinger B.V., P.O. Box 810, 2160 SZ Lisse, The Netherlands; Vols. 21-29, 31-47.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £30·00.

They are entitled to purchase any or all of the Society's Journals at the special rates of £6·00 per Journal per annum.

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, St. Andrew's House, 48 Princess Road East, Leicester LE1 7DR.

BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY,
VOL. 54, PART 3, 1984

CONTENTS

ARTICLES:

Michie, Susan — Number Understanding in Preschool Children . . 245

Wood, H. A., Wood, D. J., Kingsmill, M. C., French, J. R. W. and Howarth, S. P. — The Mathematical Achievements of Deaf Children from Different Educational Environments 254

Saunders, Robyn J. and Solman, Robert T. — The Effect of Pictures on the Acquisition of a Small Vocabulary of Similar Sight-Words 265

Seddon, G. M., Tariq, Riaz H. and Dos Santos Veiga, J. — The Transferability of Two Pictorial Scientific Tasks between Different Spatial Dimensions. . . . 276

Chapman, James W., Silva, Phil A. and Williams, Sheila M. — Academic Self-Concept: Some Developmental and Emotional Correlates in Nine-Year-Old Children . . 284

Furnham, Adrian — Getting a Job: School-Leavers' Perceptions of Employment Prospects 293

Kozéki, B. and Entwistle, N. J. — Identifying Dimensions of School Motivation in Britain and Hungary 306

Reid, K. — The Behaviour of Persistent School Absentees. . . 320

Trueman, M., Lynch, A. and Branthwaite, A. — A Factor Analytic Study of the McCarthy Scales of Children's Abilities 331

RESEARCH NOTE:

Fraser, Barry J. — Differences between Preferred and Actual Classroom Environment as Perceived by Primary Students and Teachers 336

Book Reviews 340

Publications Received 345

Author Index to Volume 54 349

1 APR 1985

UK ISSN 0007-0998